DOCUMENT RESUME

ED 427 447 EC 307 009

AUTHOR	Callahan, Carolyn M.; Tomlinson, Carol A.; Moon, Tonya R.; Tomchin, Ellen M.; Plucker, Jonathan A.
TITLE	Project START: Using a Multiple Intelligences Model in Identifying and Promoting Talent in High-Risk Students. Research Monograph 95136.
INSTITUTION	National Research Center on the Gifted and Talented, Storrs, CT.
SPONS AGENCY	Office of Special Education and Rehabilitative Services (ED), Washington, DC.
PUB DATE	1995-09-00
NOTE	145p.
CONTRACT	R206R00001
AVAILABLE FROM	NRC/GT, University of Connecticut, 362 Fairfield Road, U-7, Storrs, CT 06269-2007.
PUB TYPE	Guides - Non-Classroom (055)
EDRS PRICE	MF01/PC06 Plus Postage.
DESCRIPTORS	*Ability Identification; Academic Achievement; Classroom Observation Techniques; Cultural Differences; Curriculum Development; Enrichment Activities; *Gifted; *High Risk Students; Intervention; Low Income Groups; Mentors; Models; *Multiple Intelligences; Outcomes of Education; Outreach Programs; Program Effectiveness; Program Evaluation; *Talent; *Talent Identification; Test Reliability; Test Validity

ABSTRACT

This monograph describes Project START (Support To Affirm Rising Talent), a three-year collaborative research effort to develop and apply gifted identification procedures based on Howard Gardner's (1983) theory of multiple intelligences. Specifically, the study attempted to: (1) develop identification procedures; (2) identify high-potential primary age students from culturally diverse and/or low economic backgrounds using the multiple intelligences model; (3) investigate the reliability and validity of the identification procedures; and (4) test the efficacy of specific interventions on student achievement and attitudes about school and self. Identified students were assigned to one of three conditions: an experimental condition involving modification of classroom activities and a family outreach program; an experimental condition involving modification of classroom activities, a family outreach program, and a mentorship; and a control group. Findings of the qualitative and quantitative study are grouped into the following categories: psychometric properties of the assessment tools; achievement, attitude, and self-concept; teacher changes during the project; outcomes for students and their families; elements of the program found to be most effective; and other qualitative findings. Two appendices include a project lesson development flow chart and classroom observation protocols. (Contains approximately 225 references.) (DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 427 447

NRC
G/T

THE NATIONAL RESEARCH CENTER ON THE GIFTED AND TALENTED

The University of Connecticut
The University of Georgia
The University of Virginia
Yale University

Project START: Using a Multiple Intelligences Model in Identifying and Promoting Talent in High-Risk Students

Carolyn M. Callahan
Carol A. Tomlinson
Tonya R. Moon
Ellen M. Tomchin
Jonathan A. Plucker

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Gubbins, E.J.
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

The University of Virginia
Charlottesville, Virginia

September 1995
Research Monograph 95136

BEST COPY AVAILABLE

EC 307009

Project START: Using a Multiple Intelligences Model in Identifying and Promoting Talent in High-Risk Students

Carolyn M. Callahan
Carol A. Tomlinson
Tonya R. Moon
Ellen M. Tomchin
Jonathan A. Plucker

The University of Virginia
Charlottesville, Virginia

September 1995
Research Monograph 95136

THE NATIONAL RESEARCH CENTER ON THE GIFTED AND TALENTED

The National Research Center on the Gifted and Talented (NRC/GT) is funded under the Jacob K. Javits Gifted and Talented Students Education Act, Office of Educational Research and Improvement, United States Department of Education.

The Directorate of the NRC/GT serves as the administrative and a research unit and is located at The University of Connecticut.

The participating universities include The University of Georgia, The University of Virginia, and Yale University, as well as a research unit at The University of Connecticut.

The University of Connecticut
Dr. Joseph S. Renzulli, Director
Dr. E. Jean Gubbins, Assistant Director

The University of Connecticut
Dr. Francis X. Archambault, Associate Director

The University of Georgia
Dr. Mary M. Frasier, Associate Director

The University of Virginia
Dr. Carolyn M. Callahan, Associate Director

Yale University
Dr. Robert J. Sternberg, Associate Director

Copies of this report are available from:
NRC/GT
The University of Connecticut
362 Fairfield Road, U-7
Storrs, CT 06269-2007

Research for this report was supported under the Javits Act Program (Grant No. R206R00001) as administered by the Office of Educational Research and Improvement, U.S. Department of Education. Grantees undertaking such projects are encouraged to express freely their professional judgement. This report, therefore, does not necessarily represent positions or policies of the Government, and no official endorsement should be inferred.

Note to Readers...

All papers by The National Research Center on the Gifted and Talented may be reproduced in their entirety or in sections. All reproductions, whether in part or whole, should include the following statement:

> **Research for this report was supported under the Javits Act Program (Grant No. R206R00001) as administered by the Office of Educational Research and Improvement, U.S. Department of Education. Grantees undertaking such projects are encouraged to express freely their professional judgement. This report, therefore, does not necessarily represent positions or policies of the Government, and no official endorsement should be inferred.**
>
> **This document has been reproduced with the permission of The National Research Center on the Gifted and Talented.**

If sections of the papers are printed in other publications, please forward a copy to:

The National Research Center on the Gifted and Talented
The University of Connecticut
362 Fairfield Road, U-7
Storrs, CT 06269-2007

Please Note: Papers may not be reproduced by means of electronic media.

Project START: Using a Multiple Intelligences Model in Identifying and Promoting Talent in High-Risk Students

Carolyn M. Callahan
Carol A. Tomlinson
Tonya R. Moon
Ellen M. Tomchin
Jonathan A. Plucker

The University of Virginia
Chalottesville, Virginia

ABSTRACT

Project Support to Affirm Rising Talent (START) was a three-year collaborative research effort undertaken jointly by The University of Virginia site of The National Research Center on the Gifted and Talented (NRC/GT) and the public schools of Charlotte-Mecklenburg, North Carolina. The researchers at the NRC/GT and the school personnel agreed to develop a program based on Howard Gardner's (1983) Multiple Intelligences (MI) theory. The specific purposes of the study were to: (a) develop identification procedures based upon Howard Gardner's theory; (b) identify high-potential primary age students from culturally diverse and/or low economic backgrounds through use of Gardner's framework; (c) investigate the reliability and validity of the identification procedures; and (d) test the efficacy of specific interventions on achievement and attitudes about school and self of identified students.

Identified students were assigned to one of three conditions: (a) an experimental condition that consisted of modification of classroom activities and a family outreach program; (b) an experimental condition that consisted of modification of classroom activities, a family outreach program, and a mentorship; or (c) a control group. Assignments to the treatment groups were made according to school the child was attending. Schools had been randomly assigned the condition. Control group students were in all schools, but their teachers were not trained in the intervention procedures and they did not have a mentor and their families were not invited to participate in the family outreach program.

As the staff of the project reflected on the results and implications of this study, we found that we learned many lessons that are not necessarily those related to talent identification and development in particular, but rather to general principles of the change process in schools. First, we found that school administrators in desperation to bring about high achievement may be institutionalizing practices that mitigate against the very change they are seeking. Second, this study unequivocally proves one point: Testing MI or similar theories of multiple intelligences is difficult, and designing instruments to test MI theory or assess multiple intelligences in students involves significant problems that individuals investigating other alternative assessments or theories of intelligence may not encounter.

Project START: Using a Multiple Intelligences Model in Identifying and Promoting Talent in High-Risk Students

Carolyn M. Callahan
Carol A. Tomlinson
Tonya R. Moon
Ellen M. Tomchin
Jonathan A. Plucker

The University of Virginia
Chalottesville, Virginia

EXECUTIVE SUMMARY

No population has either a monopoly on or an absence of talent, but in many areas across the nation, minority and at-risk students are not represented in gifted programs in proportion to their presence in the general population. Students who are at-risk because they live in impoverished environments are greatly under-represented as are minorities in many communities (Baldwin, 1994; Frasier & Passow, 1994; U.S. Department of Education, 1993). While many attempts have been made to identify means of rectifying inequities, especially with the funding of model projects through the Jacob K. Javits Gifted and Talented Students Education Act of 1988, there is little systematic research on the effectiveness of alternative models of identifying these students or on effective services and intervention with these groups of learners. Further, there is little research on effective means to nurture talent among those children whose families do not have the means to provide the stimulation and experiences available to children of middle-class and upper-class environments (Patton, Prillaman, & VanTassel-Baska, 1990).

Our project was an attempt to study the effects of systematic talent identification and development among children from minority populations and children of poverty at the primary school level. We sought to address the major question: In what ways can schools successfully identify students from minority and/or at-risk populations and provide experiences that will result in the emergence of extraordinary abilities?

Research Questions

Specific research questions were grouped according to four general areas of investigation: (a) evaluation of the psychometric properties of the instruments used in the identification process, (b) student outcomes, (c) teacher impact, and (d) student impact.

Assessment Questions

1. Do the MI-based assessments show evidence of reliability and validity?
2. Are there ethnic and/or gender differences in student performance on the MI-based assessments?
3. How do the MI-based, alternative assessments compare to traditional, standardized assessments with respect to psychometric properties?

Student Outcome Questions

4. What impact do the interventions have on student achievement?
    - a. Do the interventions have different impacts on student achievement in four different academic subjects (vocabulary, reading, language arts, mathematics)?
    - b. Do the interventions have different impacts on student achievement based on student ethnicity across the four subject areas (vocabulary, reading, language arts, mathematics)?
5. What impact do the interventions have upon student attitudes toward learning and education?
    - a. Do the interventions have different impacts on student attitudes towards learning, teachers, language arts, and/or mathematics?
    - b. Do the interventions have different impacts on student attitudes based on student ethnicity toward learning, teachers, language arts, and/or mathematics?
6. Do Project START interventions have an effect on student self-concept?
    - a. Do the interventions have different impacts on specific areas of student self-concept?
    - b. Do the interventions have different impacts on student self-concept based upon student ethnicity?
7. Are students who participate in Project START interventions referred and selected for entrance into the Charlotte-Mecklenberg Schools gifted program at a higher rate than control students?

Teacher Impact Questions

8. How do teachers view the diversity of students in their classrooms after exposure to a model of individual differences such as that of MI theory?
9. How do teachers make meaning of the task of identifying student talent in diverse populations after exposure to MI theory?
10. How do teachers integrate notions of working with underserved students with a curriculum stressing language immersion, multiculturalism, manipulatives, and multiple talents and intelligences?
11. Are there specific developmental processes that teachers undergo as they attempt to implement a MI-based perspective on education, learning, and talent?

Student Case Study Questions

12. In what ways do the interventions impact the lives of individual Project START students?
13. What are the effects of encouraging the parents of underserved students to become more involved in their children's education?
14. What are the differing dynamics in the lives of children judged to be successful and unsuccessful in Project START (according to changes in achievement test scores and teacher perceptions)?

Methods

Sample

Cohort 1. The sample for this portion of the study consisted of 1,813 children enrolled in kindergarten and first-grade in 16 Project START schools at Charlotte-Mecklenburg Schools (CMS) in Charlotte, North Carolina during the 1992-1993 school year. The target population for assessing the instruments was ethnically diverse and/or low SES kindergarten and first-grade students. Female students comprised 48.2% (*n* = 873) of the sample, and ethnic composition of the students was as follows: Caucasian, 18.8%; African American, 71.3%; Asian American, 1.8%; Hispanic American, 2.5%; and other ethnic groups, 3.5%. Socioeconomic status (SES) was determined by participation in the federal government's free/reduced lunch program. The lunch status could not be determined for 59.8% of the sample because of issues of confidentiality of data for those not selected for the project. Of the remaining 729 students, 48.4% (*n* = 353) receive free or reduced lunch.

Data on additional measures of attitude, self-concept, and achievement were obtained for the 371 children who were selected for the talent development program and were randomly assigned to experimental or control conditions. In the validity sample, 49.6% (*n* = 184) of the students were female, and the ethnic composition of the students was as follows: Caucasian, 23.5%; African American, 63.6%; Asian American, 6.5%; Hispanic American, 2.7%; and other ethnic groups, 3.8%. Nearly half of the students (48.2%, *n* = 179) received free or reduced lunch.

Cohort 2. The students in cohort 2 (*n* = 1,077) attended kindergarten during the 1993-1994 school year. The sample was roughly balanced by gender (51.1% female), with the following ethnic composition: Caucasian, 27.0%; African American, 66.6%; Asian American, 1.9%; and Hispanic, 4.5%. A majority of the students was considered to be economically disadvantaged by school district personnel.

Instrumentation

Cohort 1. In the spring of 1993, the subjects were assessed using the Multiple Intelligences Assessment Technique (Udall & Passe, 1993) with local modifications. The battery of assessments consisted of 13 performance-based activities, teacher ratings, and observational checklists corresponding to four of the multiple intelligences: logical-mathematical, linguistic, spatial, and interpersonal. In order to obtain estimates of construct (discriminant and convergent) validity, the Iowa Tests of Basic Skills (ITBS; language arts, mathematics, reading comprehension, and vocabulary sub-tests) were administered.

Cohort 2. The assessment battery used with cohort 1 was modified for use with cohort 2 during spring of the 1993-1994 school year. Two observer scales were added to provide ratings of each student's use of linguistic skills during the assessment process and use of interpersonal skills during other classes such as music and physical education. One performance activity (disassembly and reassembly of a mechanical pump) was replaced with another (disassembly and reassembly of a mechanical drain). With the modifications, the battery included four classroom teacher rating scales, three observer or non-classroom teacher rating scales, and eight performance assessments.

Data Analysis

Quantitative

Cohort 1. Because of logistical problems associated with large-scale performance assessments, test-retest and inter-rater reliability could not be obtained for the performance assessments, although obtaining this evidence is a priority of future administrations of the assessments. Cronbach's alpha was calculated as a measure of internal consistency for each of the four subscales (verbal-linguistic, logical-mathematical, spatial, and interpersonal).

Correlations among the assessment subscales and ITBS subscale scores were computed to obtain evidence of construct validity. Confirmatory factor analysis was used to determine whether the activities assessed the corresponding four intelligences (verbal-linguistic, logical-mathematical, spatial, and personal). An examination of convergent and discriminant validity was conducted using multitrait-multimethod matrices.

Fixed effect analyses of variance (ANOVAs) were conducted using the ITBS subscale scores as dependent variables to determine the presence of any differences based on gender, ethnicity, or school.

Cohort 2. Internal consistency and inter-rater agreement estimates were calculated using each student's data, while a random sample of students were administered specific performance measures after approximately four weeks to allow calculation of stability estimates. Traditional (Campbell & Fiske, 1967) and structural equation modeling (Jöreskog & Sörbom, 1989) approaches to multitrait-multimethod analysis were conducted to gather evidence of construct validity. The existence of gender and ethnic differences were investigated using multivariate analysis of variance (MANOVA) techniques.

Assessment of student linguistic-pictorial performance was accomplished through the use of up to two activities: oral skills and emergent writer. Students were allowed to write, describe, or both write and describe their pictorial work in order to allow for differences in student thinking style. Therefore, reliability calculations were performed separately for the emergent writer ratings and oral skills ratings.

Achievement, Attitude, and Self-Concept

The research questions formulated for this portion of the study were answered using repeated measures ANOVAs. Separate ANOVAs were conducted for each type of outcome (e.g., ITBS, attitude surveys, and self-appraisals), and separate sets of analyses were performed for each cohort. Because of sample attrition across the project period and incomplete data, the sample for this portion of the study was substantially reduced from the previously reported sample making multivariate techniques inappropriate.

Qualitative

Data were gathered through four primary means: (a) written surveys administered at the end of the first summer institute and near the end of the three yeas of interventions, in which teachers were asked to share their understanding of various START components and procedures; (b) persistent observation in START classrooms using structured observation protocols throughout the duration of the project; (c) use of focus group interviews with district resource teachers in programs for learners identified as gifted; and (d) formal and informal interaction with participating teachers and principals throughout START.

Qualitative study, using observation and interview, allows home-school connections to be explored, enables researchers to look for indicators of success or lack of success not likely to be reflected on standardized measures, and provides a mechanism for examining the impact of varied program components (singly and together) on the lives of students and families.

Data analysis regarding teacher change was on-going throughout the duration of START, guiding both subsequent staff training and data collection. A constant comparison method of coding and theme emergence was employed with prose classroom observation records, teacher surveys, focus group notes, and researcher field notes. Member checks were used with participating teachers in two ways: (a) observations were followed by teacher/observer conferences in which observers reflected what they felt they were seeing during the observation and then discussed with the teacher their perceptions of the class, and (b) large group discussions between researchers and teacher participants provided an opportunity for researchers to share patterns that they felt they were seeing, and to ask the teachers to corroborate or modify the conclusions. Peer debriefing was employed in three ways: (a) periodic joint classroom observations conducted by the project consultant and project coordinator in which each took observational notes and then compared notes for similarities and differences following the shared observations, (b) formal sharing of patterns noted in separate observations by the project consultant and project coordinator twice yearly during the three-year project span, and (c) regular peer debriefing sessions among staff members at the NRC/GT comparing interpretations of START data for this segment of the project.

Researchers each spent a minimum of six days on-site in the students' classrooms and conducting interviews with parents. Three site visits spread over three months allowed researchers to follow up with questions raised in early analysis of data and allowed for observation of students over time.

Following each field visit, researchers paired to debrief field notes, observation notes, and transcripts. Researchers coded notes for recurrent patterns and ultimately themes. In addition, notes and transcripts were reviewed by a research coordinator to probe for additional questions, ambiguities, and themes. Each child's case was individually constructed by his/her primary researcher according to a protocol developed to promote categorical consistency across cases. Categories included a description of school and home settings, a vignette of a typical school experience, the child's involvement in START (from observation as well as perspectives of parent, child, family outreach coordinator, mentor coordinator, principal, and other key players), curricular modifications observed and/or reported, evidence of the child's talent (strength) areas, and additional themes and impressions. Group debriefing sessions were held between each site visit to look for emergent common and disparate themes among the cases.

Results

The Psychometric Properties of the Assessment Tools

Reliability evidence suggests that the subscales are internally consistent. The factor analysis confirmed the presence of the linguistic and logical-mathematical subscales, but the presence of the two remaining subscales could not be confirmed. The combination of linguistic and interpersonal intelligence activities on the first factor is not surprising, since interpersonal communication contains a major verbal component.

No meaningful gender differences were found on the assessments, but the relatively high ratings of Asian American students are cause for concern. Teacher subjectivity may be influencing the assessments, or the use of performance assessments to avoid ethnic bias on standardized tests may simply be misguided. Significant differences in ratings among schools exist (although they are associated with small effect sizes), which supports previous research on the inconsistency associated with performance-based assessments (e.g., Aschbacher, 1991; Haertel, 1994).

Analysis of the multitrait-multimethod matrix yielded limited evidence of both convergent and discriminant validity. A possible explanation for the relatively large correlations among the mathematical-logical assessments (tasks and checklists) and ITBS subtests may be that the activities and checklist associated with the math-logical subscale are the most objectively scored of those in the MI-based alternative assessment battery.

Achievement, Attitude, and Self-Concept

In summary, the results from the quantitative analysis of achievement data suggest that the intervention had no significant effect on the achievement of the treatment groups. All children in the study demonstrated growth, but there were no significant differences among groups. On attitudes toward learning, teachers, language arts and mathematics, the students in all three groups in cohort 1 experienced a decrease in attitudes toward learning and teachers. In cohort 2, attitudes toward teacher and language arts increased significantly for all three treatment groups, but attitudes toward learning also decreased for this group.

Qualitative data, on the other hand, suggested improved student attitudes about self as student and about one or more or more aspects of school—particularly mentorship and family outreach components of START. Parents often also saw their children as more successful as a result of START, believed themselves to have developed stronger parenting skills through START, and often felt more welcomed and comfortable in their child's school as a result of participation in START events. Following are key themes from the qualitative portion of the study.

How Did Teachers Change During the Span of the Project?

Establishing a START classroom called upon teachers to understand and incorporate instruction that was: concept-based, differentiated in response to learner need, centered on use of concrete manipulatives at appropriate points, sensitive to the need for consistent use of rich language by the student, and built around multiple intelligences. That is, of course, a complex charge under the best of circumstances. Virtually all participants made observable change in some of these areas. Few were able to develop proficiency in all areas. Among patterns of teacher development evident among START teachers over the three year span of the project were:

- Teachers rapidly learned language and concepts related to the project (knowledge about START goals). With relative facility, they learned to discuss ways in which they might incorporate the various project elements into their classrooms (knowledge of how to execute START goals). Translation of understanding into practice came much more slowly. By the end of the project, however, most of the first and second-grade teachers routinely incorporated one or more START elements into their instructional planning and routines.

- Use of multiple intelligence theory was powerful in two ways. First, it gave teachers an optimistic way to think about students with non-traditional strengths. If START students did not appear strong with language or math, MI provided a concrete mechanism to look for other ways the students demonstrated potential. Second, use of MI was a powerful lever for moving from more rigid and teacher-centered places to more flexible and student-centered ones. If a teacher was to present a lesson through several MI modalities, then students had to work in small groups, multiple centers or tasks would be required, and automatically it was necessary to diminish the "frontal" role of the teacher. By the middle of year 1 of the project, classrooms of START teachers observably contained more student movement, student-to-student interaction, and use of varied instructional materials than did primary classrooms of non-participating teachers. That pattern was maintained throughout the project duration.

- Growth in concept-based instruction was difficult for the teachers. Apparently due in large measure to heavy district emphasis on high stakes state tests, virtually all of the teachers' instruction was skills-based. In rare instances, a topic such as an animal or a holiday would be the focus of lessons. During START, teachers developed at least a general understanding of concept-based instruction, and worked together in cross-school teams to create concept-based "learning tubs." These large, plastic boxes contained outlines of concept-based learning experiences, materials necessary for using them in the classroom, and guidelines on teaching the lessons. The goal of the tubs was shared creation of concept-based lessons that contained varied MI emphases for student activities and that embedded skills goals in a meaning-rich context. For some teachers, use of concept-based instruction was a powerful sense-making device for themselves and their learners. Other teachers resisted extensive use of concept-based instruction, at least in part, because of fear that any departure from drill on skills would damage student scores on state tests.

- Teachers' use of differentiated instruction in response to learner need could be viewed in two ways—one more optimistic, one less so. START teachers demonstrated consistent responsiveness to student learning profile and interest differences through their use of varied MI modalities for student activities and products. They were less able to plan for student differences in readiness. In other words, a student who was extremely able spatially would work with the same spatial task as a child who liked spatial activities, but who showed little advanced skill with them. A child who was advanced with a given math skill would generally have the same math task as learners less advanced with the same skill.

- START teachers also typically became advocates for their START students as the project progressed. Three patterns were evident among teachers in this regard. First, teachers were generally willing to accept that a child was "smart" based on data from the START identification procedure. Second, teachers often became eager to play a role in making sure START students had a chance to develop their capacities not only in the START classroom, but also through the district's program for learners identified as gifted. Third, while teachers accepted that a child might, for example, have kinesthetic strengths because START identification indicated that was the case, they were not willing to accept that as "the whole truth." Throughout the study, teachers insisted that they found MI strengths in START students other than those indicated initially.

- While teacher attitudes about START students generally became more positive, there was still a tendency in START teachers to associate student "success" with student compliance. "High maintenance" students, those with demanding personalities and behavior patterns, remained more likely to be seen by their teachers as "unsuccessful" in the program. Conversely, students who either had a particular strength in the traditionally valued areas of reading or math, or who were compliant were generally perceived by their teachers to be successful in START.

- Although considerable evolution of teacher thought and practice was evident throughout the duration of START, it seems likely that project goals were complex enough to require longer than three years for mastery by teachers. Thus while evidence of positive teacher change was clear via observation, interview, and artifact examination, translation of this growth into potential student achievement, as measured by standardized tests, would likely not result until greater teacher proficiency and automaticity were evidenced.

How Did START Impact Participating Students and Their Families?

The impact of START on lives of participating students and their parents was clearly positive. In fact, the opportunity to take part in START was such a major event for participating students and parents that it spoke loudly about both the need of families under stress for positive messages, and the eagerness of such parents to participate in the growth and development of their children. Among evident themes in case studies of eight START students related to the impact of START on participating students and their parents are the following:

- Parents of all eight students defied the stereotype of unconcerned or uncaring parents. Despite barriers of language, poverty, fear, and transportation difficulty, every one of the parents studied was eager to make life better for the children. While they might not always "understand the system" or feel comfortable with it, they wanted to do whatever they could to ensure that the system worked well for their children. While they might not always know how to help their children with school, they wanted to learn how.

- Each of the START parents responded with energy and enthusiasm to the positive message sent to them by the school about their child when the child was invited to participate in START. Many of the parents remained uncertain about the meaning of multiple intelligences (in spite of continued communication from the district about the meaning of the project), but they clearly heard the message, "Your child is smart, important, and worthy of attention." For many of the parents, this translated into the conclusion, "I must be doing something right." Life is difficult for many START families, making the presence of a powerful, positive message from school both valuable and rare.

- Because START teachers were ready to accept the idea that START students were talented, even if not in "traditional" or "standard" ways, it seems likely that START students began to experience what it feels like when a teacher believes in you. Noted one teacher, "I once would have seen these children as different, with problems. Now I see them as different with potential." Most of the case study students felt the positive teacher response and reported throughout the duration of START that they liked school, and most often, that they were good at school. START also provided an expectation that START teachers communicate with parents of identified children throughout the year. It appears that generally positive teacher attitudes about START learners were communicated to parents who felt accepted by the teachers and welcomed at START school activities. Most of the START parents attended many school functions.

What Elements of START Worked for Participating Students?

All of the case study children appear to have succeeded in noteworthy ways as a result of START. Increased attendance at school, greater enthusiasm for school work, growing self-esteem, increased language skills, rising standardized test scores, and identification for participation in the district's program for learners identified as gifted are examples of positive developments for the case study students during their time in the project. Four elements of START were examined for their effectiveness with case students involved in the program.

- **Multi-Avenue Approach**—In each school, several adults and programs were focused on START learners. The classroom teacher, a family outreach coordinator, a mentorship coordinator, and sometimes other support personnel such as a second language teacher, a reading specialist, or a gifted education specialist were involved with START. School principals were typically also highly engaged in making certain the program worked. In no instance, did all the adults or all of the program components work effectively for a case study student. In every case, however, some adults and program components were powerfully effective. Having several key adults and program elements attending to the needs of the learner meant that if one element was weak, the capacity of the program to have a positive impact was not extinguished. The result of the multi-faceted approach was that each case study student had one or more adults "in his/her corner," and one or more program elements able to make a difference in his/her life.

- **Mentorship**—For all the case study students and their parents, the most overt enthusiasm was evidenced for the mentorship component of START. Although students often met with their mentors only once a week, met in the school, and experienced relatively weak linkage with the regular classroom, the mentorship remained very important. There seemed to be at least two reasons for the power of this program component. First, the mentors were often quite effective with the students—able to communicate with them and invested in doing so. Second, the mentorship singled out the START student as important in a concrete and visible way. START students left their classrooms because someone from the community came to work with them. For the students and their parents, this was again a message that the child was important.

- **Family Outreach**—Family outreach activities were planned at each school several times a year for START parents and often for the children as well. Schools often provided transportation and food. Programs provided information on parenting, offered a time for family recreation, or had START children make presentations to the parents. Case study parents reported that these events made them feel good because they could see their children shine, because parenting information was genuinely helpful, and because they felt less isolated as a result of the meetings.

- **Classroom Modifications**—In general, START interventions made classrooms more accepting of or enthusiastic about START children than would likely have been the case without START. The program also resulted in classrooms with more student movement, more student choice, and more teacher and student conversation about developing potential. In some affirmative way, START students were "stars" in their classrooms—a focus of much positive teacher attention and effort. While particular instructional components of the program required considerable time for teachers to implement effectively, the affirming nature of the environment was generally both apparent and significant to these learners.

What Other Qualitative Findings Are Useful?

At least three additional themes recurred in qualitative data. Each is of importance in developing and implementing future projects such as START.

- **Related to Use of Multiple Intelligence Theory in the Classroom**—Many materials on the market related to multiple intelligences are activity driven, fragmented, and incoherent. Few relate to escalating competencies in a given intelligence area, place in a coherent curriculum, or the nature of a discipline or intelligence area. One result is that MI is more fad-like than authentic in helping teachers and students develop talent. A second result is that MI in practice become more of an approach to learning style than to developing intelligence. A great deal more work needs to be done on how to ensure that MI is part of a rich and robust curriculum. Right now, it is easy to have MI become a shallow panacea.

- **Related to Site Selection for Research**—While the district involved with START embraced participation in the project and remained generally supportive during its duration, teachers felt severely compromised by participation in the project. They felt caught between two powerful and competing messages. First, the district was supportive of START. Considerable time, expense, and rhetoric was aimed at teacher growth in project goals. Simultaneously, there was something approaching a mania about student test scores on state standardized tests afoot in the district. Teacher salaries, job security of principals, and media focus gave them clear messages that the number one item on the district agenda was test score improvement. While START asked teachers to develop more student-centered, meaning-laden, constructivist classrooms in which student differences were honored, the district seemed to mandate a lockstep, skills-based, drill and practice approach to learning. In the end, stress on teachers was considerable. They felt they were always "letting somebody down," and became confused about what constitutes effective teaching. This paradox of positions was not readily apparent at the genesis of the project, raising the caution that a site may overtly embrace involvement in a particular initiative without realizing its other goals will make project participation and effectiveness difficult. Mutual attempts to identify conflicts in values and agreements to free participating teachers from such constraints is important at the outset of a major research intervention in classrooms.

- **Related to Project Duration**—During the three year START intervention, it appears that teacher development was evident. It also seems clear that the change was not yet of a degree to yield change in student achievement as measured by standardized tests. In addition, the change was certainly not fully institutionalized in teacher practice. Perhaps most important, student and family involvement in START was important to identified learners and their parents. For all these reasons, it is important to plan for longer funding cycles for projects focused on teacher development and family impact, and/or to seek alternative sources for extended work in sites where significant change is desired. Absent those options, both people and desired outcomes are likely to be abandoned too soon.

- **Related to the Utility of Qualitative Research**—Qualitative approaches to understanding complex school environments that include teacher, student, and parent perspectives seem important in this study. The relatively brief duration of a project focused in large measure on teacher change diminishes the likelihood of student achievement effects. Qualitative data from this study, however, would suggest that noteworthy changes in the lives of students, parents, and teachers resulted from project participation. Without qualitative approaches that probe personal meaning and look for patterns that may precede or superseded quantitative data, it seems evident that researcher awareness of intervention effects would be considerably lessened. Combined use of qualitative and quantitative methods seems well advised in attempting to understand multi-faceted and complicated educational settings.

Conclusions

Our findings support the value of combining quantitative and qualitative data analysis. The quantitative data allowed us to examine overall impact on groups of children. While we found the quantitative data quite mixed, qualitative data revealed specific program impacts on students and parents, changes in teaching practices, and the relative effectiveness of the various program components. These insights were not available through quantitative data. We concluded:

1. Educators must develop images of high-potential, high-risk learners which extend well beyond the classroom. These students and their families have complex lives which—positively or negatively—profoundly affect relationships at school and engagement in learning.

2. In considering creation of intervention with teachers that will be powerful enough to predict substantial impact on the learning of high-risk, high-potential learners, at least two themes recurred. First, if such teacher change is to occur, substantial time and consistent coaching are required.

   A second theme among teachers of case study students is the need to select sites for funding where district and/or school initiatives are in harmony with goals of the grant being funded. Project START goals included the development of student-centered classrooms in which students have an opportunity to construct meaning via high relevance content and activities, to use non-traditional assessment mechanisms, and to provide numerous avenues to learning so that children with strong talent in non-traditional areas recognize themselves and their potential through the curriculum. However, the district in this setting had a heavy emphasis on raising student test scores on traditional, skill-based tests—to the degree that teacher salaries were linked to ascending test scores, and as one teacher noted, "principals disappear overnight if the test scores don't go up." START teachers found it confusing and difficult to embrace these competing approaches simultaneously.

3. Several of the case study students were identified for and participated in the district's program for gifted learners, a major goal of Project START. Generally, support staff in the program for the gifted appear enthusiastic about working with these students, interested in their progress, and aware of adjustment issues that face these students as they enter such a program. Less evident is a specific plan for undergirding START student success in academically advanced programs and classes. There was also less evidence of differentiation of instruction in special classes for START learners (and others who might benefit from it) to address experiential gaps, learning styles, issues of content relevance, and so on. When programs such as START are created, at least in part to promote equity of identification for and participation in programs for students identified as gifted, it would seem important to emphasize success-building for those students who do participate in the services that open up to them as a result of the initiatives.

References

Aschbacher, P. R. (1991). Performance assessment: State activity, interest, and concerns. *Applied Measurement in Education, 4*(4), 275-288.

Baldwin, A. Y. (1994). The seven plus story: Developing hidden talent among students in socioeconomically disadvantaged environments. *Gifted Child Quarterly, 38*(2), 80-84.

Campbell, D. T., & Fiske, D. W. (1967). Convergent and discriminant validation by the multitrait-multimethod matrix. In D. N. Jackson & S. Messick (Eds.), *Problems in human assessment* (pp. 124-132). New York: McGraw-Hill.

Frasier, M. M., & Passow, A. H. (1994). *Toward a new paradigm for identifying talent potential* (Research Monograph 94112). Storrs, CT: University of Connecticut, The National Research Center on the Gifted and Talented.

Haertel, E. H. (1994). Theoretical and practical considerations. In T. R. Guskey (Ed.), *High stakes performance assessment: Perspectives on Kentucky's educational reform* (pp. 65-75). Thousand Oaks, CA: Corwin Press.

Jöreskeg, K. G., & Sörbom, D. (1989). LISREL7: A guide to the program applications. Chicago: SPSS, Inc.

Patton, J. M., Prillaman, D., & VanTassel-Baska, J. (1990). The nature and extent of programs for the disadvantaged/gifted in the United States and territories. *Gifted Child Quarterly, 34*(3), 94-96.

Udall, A. J., & Passe, M. (1993). *Gardner-based/performance-based assessment notebook.* Charlotte, NC: Charlotte-Mecklenburg Schools.

U.S. Department of Education (1993). *National excellence: A case for developing America's talent.* Washington, DC: Office of Educational Research and Improvement.

Table of Contents

ABSTRACT v

EXECUTIVE SUMMARY vii

CHAPTER 1: Introduction and Review of the Related Literature 1
- Introduction and Overview 1
- Research Questions 4
- Review of the Related Literature 5
  - Talent Identification and Talent Development in Underserved Populations 5
  - Talent Identification in Young At-Risk Students 6
  - Correlation of Identification and Programming 8
  - The Use of Performance Assessments to Identify Talent in Diverse and Economically Disadvantaged Populations 8
  - Multiple Intelligences Theory 10
  - MI Theory and Performance-Based Assessment 10
  - Constructivism 10
  - Teacher Change 13
    - The Costs of Change 13
    - Meeting the Costs of Change 14
    - Changes Are Driven by a Sense of Purpose 14
    - Changes Are Well Designed, Realistic, and Take Into Account the People Involved 14
    - District Staff Provide Clear, Consistent Leadership and Support for the Reforms 15
    - Principals Are Actively Involved and Provide Clear and Consistent Direction in All Phases of the Process 15
    - Participants Take Part in Changes Voluntarily 15
    - Teachers Are Provided With a Variety of Forms of Technical and Psychological Support 16
    - Teachers Work Collaboratively With Other Teachers 16
    - Teachers Take Leadership Roles 16
    - The Community and Culture of the Schools Are Taken Into Account 17
    - Leaders Understand That Change Takes Time and Is Often Met With Resistance 17
    - Conclusions 17
  - Benefits of a Multicultural Approach to the Education of Minority and Low Socioeconomic Status Students 17
    - Definitions 17
    - Curriculum Content 18
    - Multicultural Education as Early Intervention or Prevention 18
    - Cognitive Styles, Ethnicity, and Socioeconomic Status 19
    - Providing All Students a Relevant Context for Learning 20
    - Multiculturalism and Self-Esteem 21
    - Outcomes of Multicultural Education 21
    - Conclusions 22
  - Mentoring 22
    - Programs With an Academic Focus 22
    - Programs With a Social Development Focus 23
    - Benefits of Mentoring 23

Table of Contents (continued)

        - Mentorships as a Factor in Talent Development 24
        - Guiding Principles for Project START 24

CHAPTER 2: An Overview of Identification and Placement of Students and Interventions 25

- Assessment, Training, and Student Selection and Placement 25
- Nature of the Intervention 26
    - Specific Training of START Teachers 26
    - Elements of START Instruction 27
    - Family Outreach 29
    - Mentor Development 29

CHAPTER 3: Methodology and Procedures 31

- Research Questions 31
- Psychometric Properties of Alternative Assessment Battery: Cohort 1 31
    - Sample 31
        - Validity Sample 31
    - Instrumentation 32
    - Data Analysis 33
        - Reliability 33
        - Validity 33
        - Gender, Ethnic, and School Differences 33
- Psychometric Properties of Alternative Assessment Battery: Cohort 2 33
    - Sample 33
    - Instrumentation 33
    - Data Collection and Analysis 35
        - Psychometric Analyses 35
- Quantitative Student Outcomes 35
    - Research Questions 35
    - Sample 36
    - Instrumentation 36
        - Iowa Tests of Basic Skills (ITBS) 36
        - Arlin-Hills Attitude Surveys 37
        - Self-Appraisal Inventory 37
    - Data Analysis 37
- Gifted Program Participation 37
    - Research Question 37
- Teacher Change 38
    - Research Questions 38
    - Project START Teachers 38
    - Data Collection and Analysis 39
        - Collecting Data Related to Teacher Change 39
        - Analyzing Data Related to Teacher Change 40
- Case Studies of Individual Students 40
    - Research Questions 40
    - Purpose of the Case Studies 41
    - Sample 41
    - Data Collection 41

Table of Contents (continued)

CHAPTER 4: Results and Discussion 43
Research Questions 43
Psychometric Properties of Alternative Assessment Battery: Cohort 1 43
Reliability 43
Validity 43
Factor Analysis 43
Construct Validity 43
Gender, Ethnic, SES, and School Differences 45
Psychometric Properties of Alternative Assessment Battery: Cohort 2 47
Descriptive Statistics 47
Reliability 47
Validity 48
Confirmatory Factor Analysis 49
Gender and Racial Differences 51
Quantitative Student Outcomes 51
Research Questions 52
Impact on Student Achievement 52
Impact on Student Achievement by Treatment Group and Student Race 60
Impact on Student Attitudes by Treatment Group 60
Treatment Group Effects 60
Treatment and Race Effects Upon Attitudes 63
Impact on Student Self-Concept 63
Cohort 1 64
Cohort 2 65
Treatment Effects 65
Gifted Program Participation 67
Research Question 67
Nature and Degree of Change Observed in START Teachers 68
Research Questions 68
In the Beginning 68
Teacher Growth in Key Components of START 70
Growth in Multiple Intelligence Instruction 71
Growth in Concept-Based Instruction 72
Growth in Differentiation of Instruction 73
Teacher as Identifiers of Talent 74
Case Studies of "Successful" and "Unsuccessful" START Students 75
Research Questions 75
Two "Successful" Project START Students 76
Charelle 76
William 77
Two "Unsuccessful" Project START Students 79
Tev 79
Denisha 80
Impact of Project START on Students and Parents 83
What Worked for the Project START Students 84
Mentorships 84
Family Outreach 85
Classroom Modifications 85
Differences and Similarities in "Successful" vs. "Unsuccessful" Project START Students 86

Table of Contents (continued)

CHAPTER 5: General Discussions and Implications 89
- The Study 89
- The Implications of the Findings 90
  - The Psychometric Properties of the Assessment Tools 90
  - Quantitative Results 92
  - Qualitative Results 93
    - Teacher Change 93
    - Case Studies of Students 95
  - Some Final Thoughts 95

References 99

Appendices

- Appendix A: Draft Model of Developing a Project START Lesson 115
- Appendix B: START Classroom Observation Protocols 119

List of Tables

Table 1 Activities Used in the Alternative Assessments Categorized by Intelligence for Cohort 1 32

Table 2 Activities Used in the Alternative Assessments Categorized by Intelligence for Cohort 2 34

Table 3 Factor Loadings, Communalities (η^2), and Percent of Variance for Principal Factors Extraction and Varimax Rotation for Performance-Based Assessment Activities—Cohort 1 44

Table 4 Interitem Correlations Among Assessment Activities and Checklists—Cohort 1 45

Table 5 Multi-Trait Multi-Method Matrix—Cohort 1 46

Table 6 ANOVA Results for Assessment Subscales—Cohort 1 46

Table 7 Descriptive Statistics for Assessment Activities and Scales—Cohort 2 47

Table 8 Stability of the Performance Assessments 48

Table 9 Multi-Trait Multi-Method Matrix 49

Table 10 Results and Goodness-of-Fit for Confirmatory Factor Analyses 50

Table 11 Means and Discriminant Analysis F-to-Remove for Gender and Ethnicity Effects Upon MI-based Assessment Scales 51

Table 12 Means (Sds) for ITBS Subtests by Treatment Group for Grade 1 53

Table 13 ANOVA Results for Treatment Effects Upon Achievement for Grade 1 54

Table 14 Means (Sds) for ITBS Subtests by Treatment Group for Grade 2 55

Table 15 ANOVA Results for Treatment Effects on Achievement by Treatment Group for Grade 2 56

Table 16 Tukey-b Post Hoc Comparisons for Grade 2 Achievement 57

Table 17 Means (Sds) for ITBS Subtests by Treatment Group for Grade 3 58

Table 18 ANOVA Results for Treatment Effects Upon Achievement for Grade 3 59

Table 19 Tukey-b Post Hoc Comparisons for Grade 3 Achievement 60

Table 20 Means (Sds) for Arlin-Hills Attitude Surveys by Treatment Group for Cohort 1 61

Table 21 ANOVAs of Retest Effect for Arlin-Hills Attitude Surveys Upon Student Attitude for Cohort 1 62

List of Tables (continued)

Table 22 Means (Sds) for Arlin-Hills Attitude Surveys by Treatment Group for Cohort 2 62

Table 23 ANOVAs of Treatment Effect for Arlin-Hills Attitude Surveys Upon Student Attitude for Cohort 2 63

Table 24 Means (Sds) of Student Self-Concepts by Treatment Group for Cohort 1 64

Table 25 Means (Sds) of Student Self-Concept by Treatment Group for Cohort 2 64

Table 26 ANOVA Results for Treatment Upon Student Self-Concept for Cohort 1 66

Table 27 Referral and Selection Information for Project START Students Currently in Grade 3 67

Table 28 Referral and Selection Information for Project START Students Currently in Grade 2 68

List of Figures

Figure 1 Schematic Design of START Treatment Conditions 3

Figure 2 Assessment, Treatment, and Testing Schedule 36

Figure 3 Vocabulary Retest x Treatment Interaction Effect—Grade 2 Achievement 57

Figure 4 Peer Self-Concept Retest x Treatment Interaction Effect—Cohort 1 66

Project START: Using a Multiple Intelligences Model in Identifying and Promoting Talent in High-Risk Students

Carolyn M. Callahan
Carol A. Tomlinson
Tonya R. Moon
Ellen M. Tomchin
Jonathan A. Plucker

The University of Virginia
Chalottesville, Virginia

CHAPTER 1: Introduction and Review of the Related Literature

Introduction and Overview

No population has either a monopoly on or an absence of talent, but in many areas across the nation, minority and at-risk students are not represented in gifted programs in proportion to their presence in the general population. Students who are at-risk because they live in impoverished environments are greatly under-represented as are minorities in many communities (Baldwin, 1994; Frasier & Passow, 1994; U.S. Department of Education, 1993). While many attempts have been made to identify means of rectifying inequities, especially with the funding of model projects through the Jacob K. Javits Gifted and Talented Students Education Act of 1988, there is little systematic research on the effectiveness of alternative models of identifying these students or on effective services and intervention with these groups of learners. Further, there is little research on effective means to nurture talent among those children whose families do not have the means to provide the stimulation and experiences available to children of middle-class and upper-class environments (Patton, Prillaman, & VanTassel-Baska, 1990).

Our project was an attempt to study the effects of systematic talent identification and development among children from minority populations and children of poverty at the primary school level. We sought to address the major question: In what ways can schools successfully identify students from minority and/or at-risk populations and provide experiences that will result in the emergence of superior abilities? The obvious first assumption we made was that talent occurs in every cultural, social, and economic group. This assumption was based on a long history of research summarized in reviews of the research by Frasier and Passow (1994) and Frasier, Garcia, and Passow (1995) and verified in the research of Hunsaker, Frasier, Frank, Finley, and Klekotka (1995).

The second assumption we made was that early identification of potential talent is critical. The talents of young children may quickly atrophy and disappear without recognition and encouragement. We further assumed that young students (grades K-2) from minority groups and at-risk populations have not yet fallen so far behind in academic achievement that teachers are unable to recognize sparks of potential talent. By broadening teachers' conceptions of talent, we presumed that we would increase the likelihood that they would seek talent using a broader array of behavioral indicators. Our intervention program was also predicated on the brain development research that indicates that the environmental influences of stimulating, engaging instruction can have its greatest impact on the rapidly growing and developing brains of young children (Kotulak, 1996).

Fourth, we interpreted the research of Bloom (1985), Csikszentmahalyi, Rathunde, and Whalen (1993), Renzulli (1994), and Feldhusen (1995) to suggest that the development of full potential requires significant support from various constituents in the child's environment, particularly the family, the school, and the community.

We identified a school system whose demographics represented the problems of under-identifying and serving students from minority groups and impoverished environments that was willing to work collaboratively to study the effects of a talent development program. That school district was the Charlotte-Mecklenburg Schools (CMS) in Charlotte, North Carolina. In this school district, there was a large gap between minorities enrolled in the elementary gifted program (8%) and the minority population of the Charlotte-Mecklenburg school system (40%). Further, approximately 69% of all CMS African American students in kindergarten through sixth-grade were economically disadvantaged, making it an appropriate setting for testing intervention effects of the nature proposed. The University of Virginia and the Charlotte-Mecklenburg Schools developed an intervention program funded through a model project grant of the Jacobs K. Javits Gifted and Talented Students Education Act of 1988 and a complementary research study also funded by the Javits program as part of The National Research Center on the Gifted and Talented (NRC/GT).

The purpose of the intervention project implemented in the Charlotte-Mecklenburg Schools, hereafter referred to as Project START, was to identify potential talent in culturally diverse or economically disadvantaged gifted primary students, to identify intervention strategies that were most effective in increasing the academic performance of these students, and to support the teachers and parents of these students. The purpose of the NR/GT research study was to investigate the psychometric characteristics of the identification procedures and to investigate the impact of the interventions on teacher behaviors and student achievement, attitudes, and self-concept.

Identification of talent in the children targeted in Project START was based on Howard Gardner's (1983) theory of multiple intelligences. Participating teachers (in grades K-1 in 16 schools) were trained to identify and work with underserved students and their parents within the framework of that model and in conjunction with specific interventions based on family outreach, mentorships, and other curricular modifications to be outlined below. Similar programs are nearly non-existent in the United States (Patton, Prillaman, & Van Tassel-Baska, 1990).

The research component of the collaborative project was designed to: (a) investigate the reliability and validity of the talent identification instrument procedures based on a series of checklists and "hands-on" activities characterized as alternative assessments derived from MI theory; and (b) assess the impact of the instructional interventions utilizing a combination of quantitative and qualitative methodologies.

Students and teachers from 16 schools participated in the study. A review of the schools' characteristics was used to classify them as having strong or weak levels of communication of high academic standards. The schools were divided into two groups along this dimension; half of the schools were considered "strong," the other half "weak." However, multiple changes in administrative leadership over the course of the study quickly made these classifications of schools irrelevant.

Treatment conditions were characterized according to three types of interventions. Type of school was considered the first treatment variable since all students in the treatment and control groups were enrolled in one of the 16 schools. Students identified as having potential talent were labeled START students. Details of the selection process are presented in Chapter 2. They were assigned to treatment or control conditions randomly. All teachers of students in the treatment conditions had training in specific classroom strategies. In one treatment group, only classroom instruction was modified. Students in the other treatment condition were exposed to a family outreach program and a mentor program as well as the classroom instruction. A control group of students (having one or more talents identified by the selection process just as the treatment students) were assigned to non-START classrooms in the same schools. Finally, in an attempt to examine halo effects, we included a small number of students in START classrooms who had nearly, but not quite, qualified as START students. They were identified to the teachers as START students and labeled "foils" in the design. Unfortunately, the small number of such students (*n* = 36) combined with high student mobility made analysis using this group for comparison impossible. (See Figure 1: Schematic Design of Project START Treatment Conditions.)

Briefly, the interventions were:

- **Classroom strategies:** The teachers participating in Project START experimental groups were provided staff development to prepare them to modify their instruction to develop language fluency, to stress a performance-oriented environment (hands-on, active participation), to emphasize multicultural values, and to use instructional strategies that would incorporate students' areas of strength as identified using the MI conception of intelligence. (More detail is provided in Chapter 2).
- **Family outreach:** Teachers and counselors recruited the active participation of parents in the talent development process. Training was offered on characteristics of gifted students (particularly as related to the areas of identified talent: verbal-linguistic, logical-mathematical, spatial, and the personal intelligences), child advocacy, nurturing children's talents, working with teachers, and volunteerism.
- **Mentor programming:** Volunteers from the community became mentors for two treatment groups of students, accepting responsibility for twice-monthly academic tutoring and enrichment experiences.

ORIGINAL

	Curriculum Modification and Family Outreach	Mentors	School* Achievement Motivation
Experimental 1	x	x	low
Experimental 2	x		high
Experimental 3	x		low
Experimental 4	x	x	high
Control 1			low
Control 2			high
Foil 1	x	x	low
Foil 2	x		high
Foil 3	x		low
Foil 4	x	x	high
Control Foil			low
Control Foil			high

MODIFIED

	Curriculum Modification and Family Outreach	Mentors
Experimental 1	x	
Experimental 2	x	x

*Change in leadership and consequent changes in school climate made these classifications meaningless by the end of the first year. Transience in school populations made the numbers in the foil groups (very small at onset (< 5)) too small for analyses.

Figure 1. Schematic design of START treatment conditions.

- **School selection:** As noted above, 50% of participating schools were identified as "high expectation schools" based on a judgment that the school community indicated higher academic expectations. Also, as noted above, these classifications were rendered meaningless by changes in administrative leadership.

Research Questions

Specific research questions were grouped according to four general areas of investigation: (a) evaluation of the psychometric properties of the instruments used in the identification process, (b) student outcomes, (c) teacher impact, and (d) student impact.

Assessment Questions

1. Do the MI-based assessments show evidence of reliability and validity?
2. Are there ethnic and/or gender differences in student performance on the MI-based assessments?
3. How do the MI-based, alternative assessments compare to traditional, standardized assessments with respect to psychometric properties?

Student Outcome Questions

4. What impact do the interventions have on student achievement?
   a. Do the interventions have different impacts on student achievement in four different academic subjects (vocabulary, reading, language arts, mathematics)?
   b. Do the interventions have different impacts on student achievement based on student ethnicity across the four subject areas (vocabulary, reading, language arts, mathematics)?
5. What impact do the interventions have upon student attitudes toward learning and education?
   a. Do the interventions have different impacts on student attitudes towards learning, teachers, language arts, and/or mathematics?
   b. Do the interventions have different impacts on student attitudes based on student ethnicity toward learning, teachers, language arts, and/or mathematics?
6. Do Project START interventions have an effect on student self-concept?
   a. Do the interventions have different impacts on specific areas of student self-concept?
   b. Do the interventions have different impacts on student self-concept based upon student ethnicity?
7. Are students who participate in Project START interventions referred and selected for entrance into the Charlotte-Mecklenberg Schools gifted program at a higher rate than control students?

Teacher Impact Questions

8. How do teachers view the diversity of students in their classrooms after exposure to a model of individual differences such as that of MI theory?
9. How do teachers make meaning of the task of identifying student talent in diverse populations after exposure to MI theory?

10. How do teachers integrate notions of working with underserved students with a curriculum stressing language immersion, multiculturalism, manipulatives, and multiple talents and intelligences?
11. Are there specific developmental processes that teachers undergo as they attempt to implement a MI-based perspective on education, learning, and talent?

Student Case Study Questions

12. In what ways do the interventions impact the lives of individual Project START students?
13. What are the effects of encouraging the parents of underserved students to become more involved in their children's education?
14. What are the differing dynamics in the lives of children judged to be successful and unsuccessful in Project START (according to changes in achievement test scores and teacher perceptions)?

Review of the Related Literature

The review of the literature is divided into seven areas, each providing background or a rationale for some aspect of the project's implementation (i.e., choice of target group, selection of identification process, intervention/treatment options) or the interpretation of the findings. The topics covered include the need to identify and develop talent in underserved populations, talent identification in young students, performance assessments, Multiple Intelligences Theory, teacher change, constructivism, multicultural education, and mentoring.

Talent Identification and Talent Development in Underserved Populations

At the core of the Jacob K. Javits Act is the assumption that there is a critical need to identify and serve students who are typically under-represented in programs for the gifted and that these students are present in all groups of students regardless of race or economic status (Frasier, Garcia, & Passow, 1995; U.S. Department of Education, 1993). The under-representation of the segments of the population comprised of minority students and students from impoverished environments has been a continuing concern of educators for several decades (Baldwin, 1977, 1984; Frasier, 1989; Renzulli, 1973). Of the many explanations offered as the reasons for under-representation of these groups, most relate in one way or another to over-reliance on traditional tests of intelligence, achievement, and teacher checklists in the identification process (Ford, 1994a; Frasier & Passow, 1994). As critics of standardized tests point out, children from the groups of concern in this study are unlikely to have had the experiences and opportunities to learn the skills, knowledge, and understandings that are assessed by traditional intelligence and achievement tests. Teacher rating scales have also been criticized as biased toward the teachers' preconceived notion of giftedness, that often is akin to high achievement in traditional disciplines through White, middle class modes of expression (Gear, 1976; Richert, 1997). The conclusion of many reviews of the issues surrounding under-representation is that there must be a new paradigm and new approaches to identification to overcome the barriers faced by the minority or economically disadvantaged child (Frasier & Passow, 1994). Hence, in this study we attempted to explore an alternative to traditional definitions of giftedness, alternative identification procedures, and interventions designed to maximize the potential of the students.

Talent Identification in Young At-Risk Students

Shaklee (1992) contends that early identification and appropriate programming are the only means through which talent can be developed in students whose environments may not provide the experiential base for maximum development of potential. Unfortunately, most gifted and talented identification systems and programs do not focus on young children. Identification usually occurs around nine years of age, due to the fact that it is easier to identify the functionally gifted child based on achievement at that time. Performance becomes more stable as children age, and there are also more reliable standardized instruments available to assist in the identification process for this age group (Karnes & Johnson, 1986).

However, the literature on the development of cognitive functioning and achievement, as measured by standardized assessments, suggests that much of the differential in performance between minority and non-minority children and between socioeconomic groups develops in the early years (Plewis, 1991). Further, the literature provides data relevant to the performance of young gifted children and guidelines for identifying the talented child (e.g., see Callahan, Tomlinson, & Pizzat, 1994) that have evolved from programs and research on these students.

Kanevsky's (1990) study of gifted and average four to eight year old children showed that gifted children (a) demonstrated greater learning, (b) had superior ability to generalize knowledge, (c) understood tasks, (d) learned from errors, and (e) recognized similarities across tasks more than the average children. The gifted children in her sample also were characterized as having greater interest in complex knowledge and independent learning. These results tend to imply that gifted children need a different learning environment than their average peers, and should be identified earlier to match their capabilities, learning needs, and learning styles to instruction.

Barbour (1992) approached the idea of early childhood gifted education by presenting what was considered the "best" in both gifted and early childhood education to design a program that would be suitable to the intersection of the two populations. In this model, assessment is viewed as the basis for programming and should be carried out in a developmentally appropriate manner (i.e., depending primarily on observational methods). According to Barbour, planning for instruction should be based on individual as well as group needs or interests. The curriculum should be both integrated and project oriented.

The National Association for the Education of Young Children (NAEYC) reinforces the notion that "the purpose of testing must be to improve services for children and ensure that children benefit from their educational experiences" (NAEYC, 1988, p. 44). NAEYC presents several basic guidelines that should be followed when assessing young children: (a) standardized tests used in early childhood programs must be reliable, and (b) multiple sources of information should be used for major impact decisions such as placement and enrollment. Tests which are chosen should be used only for the purpose for which they are intended and for which data exist demonstrating validity. The tests results should be interpreted only by those knowledgeable about testing, as well as interpreted accurately and cautiously to parents, personnel, and the media. Tests used to assess achievement or to evaluate a program should be consistent with the goals and theories, philosophies, and/or objectives of the program. The test administrator must also be aware of individual diversity, which would include developmental as well as racial, and socioeconomic differences.

An additional tenet which Frasier and colleagues hold as important in identifying under-served gifted students includes recognition that the gifted are not a homogeneous group nor do they all express talents the same away (Frasier, Garcia, & Passow, 1995;

Frasier & Passow, 1994). Gifted students may be gifted along only one dimension of performance (e.g., language usage or verbal ability, mathematical ability, spatial ability, dance) or along several. Further, not all verbally talented children will use written, standard English to express their talents. For example, some children may be from a culture steeped in the oral storytelling tradition; some may express their abilities in rap lyrics.

Finally, the NAEYC (1988), Renzulli (1986), and Abeel, Callahan, and Hunsaker (1994) concur with many other experts in the field of assessment of young children that a shift in "emphasis from absolute concept of 'being gifted' (or not gifted) to a concern with developing gifted behaviors in those youngsters who have the highest potential for benefiting from special educational services" (Renzulli, 1986, p. 61). Hence, data collected during the identification process should be used to help determine the curriculum of the program.

Formal data include scores on intelligence tests and other norm- and criterion-referenced measures. Informal sources of data tend to be qualitative in nature and include observations and judgments using checklists or ratings of creative and productive thinking skills, creative use of words, ability to solve problems, leadership skills, and skills in the visual and performing arts. Informal sources may also include anecdotal records; teacher, parent, peer, community nominations; portfolios and/or products of the child, and interview data (Karnes & Johnson, 1986).

Richert (1987) suggests six alternative principles of identification of economically disadvantaged students:

1. Pluralism: The broadest defensible definition of giftedness should be used to include a diversity of abilities as evidenced by different populations.
2. Defensibility: Procedures should be based on the best available research and conceptualizations of the nature and needs of the gifted.
3. Comprehensiveness: As many students as possible with potential should be identified and served.
4. Advocacy: Identification and programs should be designed to meet students' needs not the preconceptions of the programmers.
5. Equity: Procedures should guarantee that no one is overlooked.
6. Pragmatism: Existing or cost effective resources should be used.

Borland and Wright (1994) cite several parallel principles to those enumerated by Richert that served as working principles enacted in this project's attempt to identify young, potentially gifted, economically disadvantaged students. First, the potential for academic giftedness is present in roughly equal proportions in all groups in our society. We predicated the START identification procedures and training of teachers on this assumption and re-iterated it throughout the project training phases. Second, identifying economically disadvantaged, potentially gifted students differs from identifying other gifted students with respect to the goal of identification. In particular, recognition and development of talent potential is a dominant factor in the process. Third, the knowledge and information needed to identify disadvantaged gifted students is to be found in school classrooms. Teachers have valuable information and classroom response to good instruction is one strong indicator of talent. Fourth, the human being is the identification instrument of choice. While tests can provide a snapshot, teacher observations can provide a story of the child's performance. And fifth, the concept of "best performance," rather than "typical" performance, is valid in identifying giftedness in young economically disadvantaged children. This is especially true for young children with minimal intellectual and experiential stimulation.

Correlation of Identification and Programming

"Best practice" in gifted education is predicated on the premise that curriculum and identification practices and procedures should be inter-related and have shared goals. Data collected as part of the identification process should provide assessment information for curricular decision-making (Kitano, 1986). Extending this logic, the activities designed to address the strengths, talents, and potential of gifted students should be related to the underlying structure of intelligence that served as a basis for identification. Hence, in Project START curricular materials and staff development we included components that reflected a model of intelligence used as the basis for identification—Gardner's Multiple Intelligences (1983) conceptions of intellect.

The Use of Performance Assessments to Identify Talent in Diverse and Economically Disadvantaged Populations

The principles discussed above regarding talent identification in young at-risk students suggest strongly that the relatively low proportion of such students in gifted programs in part may be a result of commonly used assessment procedures. In particular, traditional standardized tests fail to identify students from this population in the numbers one might expect because these tests contain cultural and/or experiential biases (Frasier, 1987). Hence, we sought to employ new conceptions of intelligence and alternative assessments to identify children who were minorities or who came from impoverished environments.

Discussions of the considerable difficulty encountered in identifying talented minority and economically disadvantaged children and the differential effects of special programs based upon student ethnicity pervade the literature (Ascher, 1988; Baldwin, 1987; High & Udall, 1983; Maker & Schiever, 1989; Ogbu, 1988). Theorists suggest that assessment procedures relying on broader definitions of "talent" and "intelligence" and the use of alternative assessment strategies will assist school personnel in more accurate identification of exceptional student abilities (Gardner, 1988; Sternberg, 1988) as well as increase representation of children from ethnically diverse populations in programs for the gifted (Patton, 1992).

Assessment terminology can be confusing and often contradictory (e.g., compare classifications used by Meyer, 1992, and Puckett & Black, 1994). For the purposes of this report, "alternative assessment" is used as the encompassing term for measures of student achievement and ability that do not employ the traditional fixed-response format typical of standardized tests. Alternative assessments in the identification of talent include such strategies as teacher and expert ratings of products and performances, peer nominations, teacher or parent ratings on observational scales, and portfolio assessment. (Note that these are not mutually exclusive categories.) Recent use of the term "authentic assessment" is predicated on assumptions made by some that fixed-response tests do not provide valid estimates of student ability and achievement.

Increased interest in alternative assessments as more valid measures of student achievement and ability is currently characteristic of the literature across educational disciplines, including science education (Collins, 1993; Doran, Boorman, Chan, & Hejaily, 1993; Finson & Beaver, 1994; Lawrenz, 1992; Reichel, 1994), the education of young children (Hills, 1992; Puckett & Black, 1994), mathematics education (Lane, 1993; Lane, Stone, Ankenmann, & Liu, 1994), reading (Peers, 1993), teacher education (Collins, 1993; Delandshere & Petrosky, 1994; Smith, Miller, & Joy, 1988), special education (Greenwood, 1994; Poteet, Choate, & Stewart, 1993; Rivera, 1993), creativity (Amabile, 1982; Baer, 1994; Hennessey & Amabile, 1988), education of the gifted and talented (Baldwin, 1994; Clasen, Middleton, & Connell, 1994), and educational measurement (Candell & Ercikan, 1994; Reckase, 1993; Shavelson, Baxter, & Pine, 1991; Stiggins & Plake, 1991).

The use of performance assessments has been quite widespread in both the United States and abroad (Desforges, Holden, & Hughes, 1994; Maeroff, 1991; Nuttall, 1992; Semple, 1992; West, Sammons Hailes, & Nuttall, 1994). A 1990 study by the Center for Research on Evaluation, Standards, and Student Testing (CRESST) found that 23 states were using, developing, or considering the use of performance assessment at a state-wide level (Aschbacher, 1991).

However, performance assessment techniques are not universally recommended. Logistical issues such as increased cost, the need for intensive teacher/observer training, more involved scoring, reliability, balance of breadth and depth of coverage, and collection of assessment materials are frequently mentioned in discussions of potential problems with performance-based assessment (e.g., Aschbacher, 1991; Darling-Hammond, 1994; Frechtling, 1991; Guskey, 1994; Marzano, 1994; Miller & Legg, 1993; O'Neil, 1992; Stiggins, 1991; Wolf, LeMahieu, & Eresh, 1992).

Psychometric issues are also raised during consideration of the quality of alternative assessments, specifically performance assessments (Burger & Burger, 1994; Dunbar, Koretz, & Hoover, 1991; Haertel, 1994; Herman & Winters, 1994; Linn, 1994; Linn & Burton, 1994; MacGinitie, 1993; Messick, 1994; Miller & Legg, 1993). As Worthen (1993) stated:

> [S]ome evidence that the technical quality of the assessment is good enough to yield a truthful picture of student abilities is essential. To succeed, alternative assessment must show that its tasks and measures are authentic (not merely authentic-looking) assessment. Otherwise, the promise it holds for improving teaching and learning will go unfulfilled. (p. 448)

Several authors question whether traditional standards of reliability and validity should be applied to alternative assessments, and if so, what types of reliability and validity should be considered (Baker, O'Neil, & Linn, 1993; Linn, Baker, & Dunbar, 1991; Swezey, 1981; Wolf, Bixby, Glenn, & Gardner, 1991). Wiggins (1993) expresses concern over the use of content, concurrent, and construct validity to demonstrate validity of performance assessments and suggests that the authenticity and face validity of the task are most worthy of educators' concern. Linn, Baker, and Dunbar (1991) stress the importance of transfer and generalizability of assessment results, while noting along with Dunbar, Koretz, and Hoover (1991), the difficulties encountered when generalizing across performance tasks. Other recommendations for establishing evidence of validity include reliance upon criterion validity and normative data (Burger & Burger, 1994) and content, concurrent, and predictive validity (Swezey, 1981). Baker, O'Neil, and Linn (1993) note that certain generalizability and reliability issues are less important when assessments are used at the group level, and similarly, Haertel (1994) recommends that, until adequate reliability and validity are assured for performance assessments, results not be interpreted at the individual level.

Few how-to and/or pro-performance assessment publications mention psychometric concerns (see Herman, Aschbacher, & Winters, 1992 and Moon, 1993 for notable exceptions). To what extent are these technical issues the concern of practitioners? Robinson (1994) believes that psychometric concerns regarding alternative assessments "are to be battled over by measurement experts" (p. 22). This attitude is considered by some to be potentially dangerous, since "internal self-criticism is rather scarce among proponents of alternative assessment. . . . The more broadly accepted [alternative assessment] becomes, the less frequently it will be challenged" (Worthen, 1993, p. 447).

Lack of critical examination may slow the development of quality alternative assessments and lessen their long-term impact upon education. Additional research into the psychometric qualities and proper use of performance assessment techniques is needed in order to guide their proper use (Baker, Aschbacher, Niemi, & Sato, 1992; Hambleton & Murphy, 1992; Miller & Legg, 1993; Stiggins, 1991; Worthen, 1993).

Multiple Intelligences Theory

The increase in emphasis upon alternative assessment techniques has coincided with the popularity of Multiple Intelligences (MI) theory (Gardner, 1983, 1993b). Current understandings of intellectual development and extensive observation of brain-damaged individuals and other special populations (e.g., autistic children, children with learning disabilities) caused Gardner to become dissatisfied with unitary models and measures of intelligence and to develop a new theory of intellectual abilities (Gardner, 1993a). The resulting theory posits seven "intelligences," although they may also be referred to as talents or abilities (Walters & Gardner, 1986). These are: linguistic; logical-mathematical; spatial, which involves knowledge and manipulation of physical things; musical; bodily-kinesthetic, the use of the body to solve problems and accomplish tasks; interpersonal, the ability to understand and interact with other people; and interpersonal, or the ability to understand one's self and use that knowledge in one's life (see Walters & Gardner, 1985, for an overview of MI theory). Although the linguistic and logical-mathematical intelligences have been historically accented in western societies and, therefore, systems of education, Gardner (1993a) believes all seven intelligences to be of equal importance and relative worth. At the time this project was conceived, Gardner had not presented naturalistic and existential intelligences.

However, MI theory is not without its critics. Several intelligence specialists have cautioned educators that MI theory still has relatively little research support (Matthews, 1988; Sternberg, 1984; Weinberg, 1989).

MI Theory and Performance-Based Assessment

The guidelines for identification that suggest alternative conceptions of giftedness as a basis for discovering talent led to the consideration of the Multiple Intelligences Model to guide the project activities. This model provides for consideration of both traditional (verbal/linguistic and logical/mathematical intelligences) and non-traditional (e.g., spatial) dimensions of intelligence. Further, Gardner (1991a) and his colleagues at Project Spectrum (Adams & Feldman, 1993; Hatch & Gardner, 1990; Kornhaber & Gardner, 1993; Krechevsky, 1991; Walters, 1992) strongly support the use of alternative assessment techniques, especially those involving performance-based assessment, to identify and evaluate student abilities with respect to MI theory. Maker (1993) and her colleagues at Project Discover (Maker, Nielson, & Rogers, 1994) are developing a series of MI-based performance assessments, and similar projects are underway across the country (see Gardner, 1993b; Maker, Nielson, & Rogers, 1994). Published research involving these efforts is not plentiful and has involved small sample sizes (Gardner & Hatch, 1989). For example, Gardner and Hatch (1989) reported distinct profiles of ability based on the Project Zero battery of assessments, but these studies involved very few students with primarily only descriptive statistics reported.

Constructivism

While MI theory provides an alternative conception of intelligence and direction for alternative assessment and strength/talent identification, it is neither a curricular or instructional model or philosophy. Accordingly, we looked to constructivism as the guide to the curriculum in the classroom.

Constructivist theory grows from the notion that meaningful learning occurs when students create new ideas from available information through reflection, experimentation, discussions, research, invention, and discovery (Bednar, Cunningham, Duffy, & Perry, 1991). Piaget believed that people go through such a process of creating new knowledge in order to solve problems that are generated by disequilibrium. Bednar, Cunningham, Duffy, and Perry (1991) and Duffy (1993) take the concept one step further to suggest that every piece of knowledge is constructed by the individual in society, no matter what the reason. Unlike the objectivist view of learning, the radical constructivist view holds that there is no objective reality (Bruner, 1986). "Reality" is simply an interaction between the individual and the surrounding social environment. (Therefore, reality is different for every society). In a sense, each interaction between the individuals of society and the environment is a creation of reality, the repeated construction of reality (Bednar et al., 1991; Bruner, 1986).

In constructivist theory, every piece of knowledge is constructed by the individual, and every creation of knowledge by the individual is an action because it "brings forth a world" (Maturana & Varela, 1987). In essence, in pure constructivist form, there is no distinction between knowledge and action. Consistent with Piaget's beliefs about human desire to eliminate disequilibrium, all action and all behavior are results of the individual trying to retain his or her organization (Cunningham, 1992; Jonassen, 1990). The interaction between the individual and the environment triggers behavior, but it does not cause it. Learning emerges from this interaction, not from the efforts to control and manipulate the environment.

Knuth and Cunningham (1993) suggest that people create worlds in order to learn and that this notion is embedded in what they call "structural coupling." Structural coupling is a more specific version of the idea that reality is the interaction of the individual and the environment. An example of such an "action" is language and the exchange of dialogue. So powerful an action is language that it is present in some sense in every human interaction, even those where actual dialogue does not take place (Cunningham, 1992; Jonassen, 1990).

Through these powerful interactions, individuals may take in multiple realities (Knuth & Cunningham, 1993). In fact, the major aim of constructivist learning is to bring forth realities or worlds that reflect the views of others, or that at least accommodate the views of others. Through dialogues with others, people can integrate multiple meaningful perspectives. Although extreme views may never be reconciled, they may at least be recognized and perhaps accepted and respected.

Simons (1993) cites six core characteristics of constructivist learning. He states that constructivist learning is "an active, constructive, cumulative, goal-directed" process that is also diagnostic and reflective. The process is "active" because students must "create" knowledge as they process incoming information, "constructive" because students must elaborate and relate complex information to stored knowledge. Constructivist learning is "cumulative" because all new knowledge that students "create" builds on prior knowledge, and "goal-directed" because it is more likely to be successful if students are aware of and help to create specific learning expectations. In addition, constructivist learning is diagnostic and reflective in that to be successful, students must engage in self-monitoring and self-evaluation techniques. In addition, Simons suggests learning should be discovery-oriented, contextual, problem-oriented, case-based, socially motivated, and intrinsically motivated.

There are several conditions that foster constructivist learning. According to Bednar et al. (1991), there should be a cognitive apprenticeship between the teacher and the student. The teacher should use a cognitive process model with the students. Over time, students assume responsibility for executing the cognitive process. Second, students should be given problems and tasks that are embedded in realistic contexts. Too often, students learn to solve "school house" problems, but can not generalize the problem-solving skills to "real life" problems. Presenting students with learning tasks that are grounded in reality serves to eliminate this obstacle.

In a constructivist learning situation, the teacher creates the environment, supplies the materials, and allows the students to identify and solve problems. The teacher serves mainly as a guide for the students. Bruner's "Man: A Course of Study" curriculum is one where this particular principle of constructivist learning is enacted (Bednar et al., 1991). Students study subjects that are regarded as important to human social behavior, such as group relationships, theories of knowledge, and culture.

A third condition that fosters constructivist learning is that of students viewing issues from multiple perspectives. Presenting students with multi-faceted problems requires them to "stretch" their own knowledge base to include perspectives of others who may not agree with them (Bednar et al., 1991; Cunningham, 1992; Jonassen, 1990). In addition, presenting real-world, multifaceted issues through group discussions where students may express their ideas allows students to discover the interconnectedness of ideas.

One important and often overlooked aspect of constructivist learning is that it is largely self-regulated (Bednar et al., 1991; Duffy & Jonassen, 1992; Jonassen, 1991). Because people see the world from different perspectives, they also learn differently. Because constructive learning relies on individual memory representations, the learners themselves play a crucial role in the learning process. No outsider can have access to the processes occurring inside students' heads. No teacher can actually make students motivated or make students engage in goal-setting or self-monitoring behavior. The role of the teacher is to create a fertile environment in which the principles of constructivist learning may be enacted. The rest is largely up to the student.

Research in the area of self-regulated constructive learning yields varied results (Bednar et al., 1991). Research shows that people view learning differently. Some believe that learning is the reproduction of ideas and facts, whereas others believe that learning is the construction of knowledge and ideas. Different learning views correlate strongly with learning styles, but correlations between learning styles and learning outcomes are in many cases insignificant (Bednar et al., 1991; Cunningham, 1992). Overall, however, constructive learners out-perform reproductive learners on surface learning, deep learning, and concrete learning tasks (Bednar et al., 1991). Because reproductive learners are strongly embedded in their learning styles, they are often intimidated by constructive learning techniques because they are very different from what is familiar to them. As a result, they distort constructive learning assignments into reproductive ones (Bednar et al., 1991).

Most of the research in constructivist learning centers around the fields of math, science, and computer science. The prevailing conclusion is that by depriving students of real-world examples, we are detracting from their learning experiences. Constructivist learning techniques may help to lend a meaningful context for learning and a strong cognitive apprenticeship between the teachers and the students (Doran & Sentman, 1994; Jonassen, 1991; Linder, 1993; Papert & Harel, 1990; Prather, 1993).

The main impediments to the practice of constructivist theory are teacher mindsets, classroom constraints, and inservice limitations (Riley, 1993). These reasons explain why constructivism is the current trend in educational research, but rarely used in the everyday classroom (Riley, 1993).

Teacher Change

The introduction of an unfamiliar conception of intellectual ability and teaching philosophy, along with expectations that teachers will understand and respond to these conceptions with a change in their behavior, presents an enormous challenge to public school classrooms. It is not a simple task to bring about change in either teachers' methods of instruction or their attitudes in the classroom. Change is difficult and complex. Fullan (1985, 1993a, 1993b) cautions administrators implementing schoolwide or systemwide changes to recognize that change is a dynamic process. Changes occur slowly and are unpredictable. A sense of realism and sensitivity to the gradual nature of change is essential to success in implementing new programs and practices.

The Costs of Change

Recent studies suggest that teachers tend to approach change slowly and hesitantly (Wallace, 1991). Most researchers agree that teachers will be reluctant to embark on the complex and difficult road to change if they are not persuaded that the benefits outweigh the costs (Cuban, 1993; Fullan, 1991; Huberman, 1983). ". . . [P]ersonal costs in time, energy, and threat to sense of adequacy, with no evidence of benefit in return, seem to have constituted the major costs of changes in education over the past 30 years" (Fullan, 1991, p. 129). A primary motivation for taking on extra work is a teacher's belief that doing so will benefit students and make him or her a better teacher (Berman & McLaughlin, 1979; Cuban, 1993). Fullan (1991) delineates the criteria by which teachers balance costs and benefits of change:

> 1. Does the change potentially address a need? Will students be interested? they learn? Is there evidence the change works, i.e., that it produces results?
> 2. How clear is the change in terms of what the teacher will have to do?
> 3. How will it affect the teacher personally in terms of time, energy, new skills, sense of excitement and competence, and interference with existing priorities?
> 4. How rewarding will the experience be in terms of interactions with peers or others? (pp. 127-128)

Existing practices and conditions within school systems often mitigate against change. If the balance is such that the costs of change outweigh the benefits, change will not occur. Among the impediments to change are inadequate allocation of time and resources for planning and skills acquisition, lack of central office support and follow through, underfunding, failure to develop school leadership, failure to recognize conflicts between proposed and existing projects, failure to account for differences between schools, lack of clarification of roles of various participants, lack of technical assistance and staff development, lack of incentives, and their presence of overall school or systemwide structures that are not conducive to instructional changes (Cuban, 1993; Pink, 1986).

Many reform efforts are impractical and unsuccessful because they fail to give appropriate attention to the factors and variables already impinging on the teachers, administrators, and classrooms in which change is expected. As Hargreaves and Fullan (1992a) point out:

> Many attempts to improve instruction take little account of the social contexts in which learning and teaching take place. The price of ignoring the context of teaching is failed idealism, guilt and frustration at not being able to meet the

> standards, criticism of teachers who fail to make the changes, and erratic leaping from one innovation bandwagon to another. (p. 56)

Meeting the Costs of Change

It is not enough to simply mandate changes and expect them to happen. There is widespread agreement in the literature on teacher change that new projects must be realistic, well developed, and substantially supported both materially and psychologically. Certain features of successful innovations are consistently described in the literature:

1. changes are driven by a sense of purpose;
2. changes are well designed, realistic, and take into account the people involved;
3. district staff provide clear, consistent, leadership and support for the reforms;
4. principals are actively involved and provide clear and consistent direction in all phases of the process;
5. participants take part in changes voluntarily;
6. teachers are provided with a variety of forms of technical and psychological support;
7. teachers work collaboratively with other teachers;
8. teachers take leadership roles; and
9. the community and culture of the schools are accounted for, and understand that change takes time and is often met with resistance (Cuban, 1993; Darling-Hammond, 1993; Fullan, 1985; Hargreaves & Fullan, 1992b; Huberman, 1983; Jackson, 1993; Sarason, 1990; Showers, 1985; Wallace, 1991).

Changes Are Driven by a Sense of Purpose

Change is more likely to succeed if it is informed by solid and substantiated theories and involves guiding principles and concepts that people involved in the change can understand (Fullan, 1985; Sarason, 1990). "The most fundamental breakthrough occurs when people can cognitively understand the underlying question and rationale with respect to 'why this new way works better' " (Fullan, 1985, p. 396). The importance of a sense of moral purpose in motivating change cannot be over emphasized, but at the same time, premature vision endangers projects (Fullan, 1993a). Emphasis on a formal, articulate vision and strategic plan in the initial reform phases is likely to represent imposition of a command from one or a small number of leaders. Moreover, premature vision and planning can be blinding. Fullan prefers building a shared sense of vision, which, having developed collaboratively over time, will better represent the true needs and experiences of the community. Change, he maintains, is a journey, not a blueprint. The process is unpredictable and problems are inevitable (Fullan, 1993a).

Changes Are Well Designed, Realistic, and Take Into Account the People Involved

Changes must be well designed and technically challenging (Huberman, 1981). It is critical that proposed changes be a good fit with the schools involved. Change involves curriculum materials, instructional practices, and teacher beliefs and understanding. Reforms must take into account and plan for the real world contexts within which changes are introduced (Hargreaves & Fullan, 1992b). If policy makers are realistic about classroom and school structures and conditions, changes are more likely to be successful and long lasting. Change agents need to be aware of the limitations within which teachers work, and account for the increased demands in terms of energy, time, and resources that even small changes in instructional practice place on personnel (Cuban, 1993).

The importance of respecting and addressing the attitudes, needs, and interests of participants and of understanding the climate and culture in which changes are introduced should not be underestimated. In addition, teachers are more likely to develop a mindset conducive to change if they feel they are in a safe environment. They are unlikely to seek help or take the risks associated with change in an environment in which failures may have negative career consequences (Showers, 1985).

District Staff Provide Clear, Consistent Leadership and Support for the Reforms

If innovations are to be long lasting, school district staff should remain actively involved throughout the change process rather than advocating or initiating change and stepping back. Researchers present consistent findings that strong, active leadership and support improve the chances that changes will occur and endure. Central staff need to be actively involved in a project throughout the process, because they can have a strong influence on a school's likelihood of follow through on proposals for change (Berman & McLaughlin, 1979; Fullan, 1991; Huberman, 1983; Pink, 1986).

Central offices can support a program in its formative phases by offering workshops, training sessions, and other forms of assistance. District level personnel are a valuable resource for teachers seeking advice and support as they implement programs in their classrooms. But their role may be even more important in maintaining than in implementing changes. They are crucial in making the transition from innovation to institutionalization. They can have a stabilizing role in the face of staff turnover in individual schools. Central office staff provide consistency by training new personnel, maintaining the project through times of financial and political crises, and incorporating new practices into curriculum and job descriptions (Berman & McLaughlin, 1979; Fullan, 1991).

Principals Are Actively Involved and Provide Clear and Consistent Direction in All Phases of the Process

Throughout the literature, no theme receives more emphasis than the critical role of the principal in the success or failure of change. Unfortunately, principals all too often ask for changes, provide some resources, and subsequently step out of the change process. While there is strong agreement that the commitment of the principal is a key aspect of change (e.g., Cuban, 1993; Fullan, 1985, 1991), Huberman (1983) differs from other authors in his added emphasis on the role of the principal in "pressuring" teachers rather than merely supporting them through encouragement and collaboration.

The active involvement of the school principal in the change process sends a message to staff about the importance of the program. In addition, the principal, as well as central office personnel, can provide consistency and support for a program in the face of staff turnover and external political pressures (Berman & McLaughlin, 1979).

Participants Take Part in Changes Voluntarily

Voluntary changes are more successful and require less support than mandated ones (Cuban, 1993; Fullan, 1985). Rather than mandate changes, it is far better to "make the innovation as attractive as possible by stressing the resources for assistance and collaboration among users" (Fullan, 1985, p. 417). In other words, potential participants need help in realizing that the benefits outweigh the costs.

For such a strategy to succeed, reforms aimed at altering teacher routines need to secure the teachers' commitment. Teachers need to be persuaded that a change will

> be better for the children, that it will not undercut their authority, and that it can be adapted to the particular setting. (Cuban, 1993, p. 281)

Teachers Are Provided With a Variety of Forms of Technical and Psychological Support

Reforms often fail because teachers are expected to add new programs and procedures to an already crowded and emotionally demanding workload without provision for the increased burdens. Teachers need to be eased into a new program with ample technical and emotional support. Release time and the use of aides and substitutes can be used to provide teachers with more time to plan, collaborate with other teachers, handle increased work loads, and make other accommodations to change (Fullan, 1985; Jackson, 1993). Support, training, and ready access to information are essential to the survival of innovations (Cuban, 1993; Darling-Hammond, 1993; Fullan, 1985; Jackson, 1993).

Teachers Work Collaboratively With Other Teachers

The importance of a collaborative relationship among teachers is emphasized in the literature on change. Change correlates positively with the extent to which teachers collaborate and lend each other technical support. Peer coaching and collaboration can provide teachers with opportunities to develop ideas, to offer one another suggestions and encouragement, and to acquire new skills without fear of evaluation (Showers, 1985).

Collaboration does not always work; it is personal and individual. Successful collaborations occur between colleagues who are able to exchange ideas within an environment of mutual trust, and who have the quality of humility required to seek and accept help (Wallace, 1991).

Collaborative skills and relationships are necessary to learning, but Fullan (1993a) reminds policy makers of the importance of individualism and dissent:

> Collaboration is one of the most misunderstood concepts in the change business. It is not automatically a good thing: it does not mean consensus; it does not mean major disagreements are verboten; it does not mean that the individual should go along with the crowd. (p. 82)

Collaboration, at its best, provides a challenging, but supportive network of collegial peers rather than the expectation of implementation of difficult innovations in isolation.

Teachers Take Leadership Roles

Studies by Berman and McLaughlin (1979) demonstrate a strong correlation between teacher participation in decision making and the success of changes. It is the teacher who is in the best position to make suggestions and offer remedies to problems that arise in practice. Teachers who act as leaders both within and outside of their classroom are more likely to implement innovations (Cuban, 1993; Darling-Hammond, 1993). This is not to imply that teachers should control the change process, but rather that leadership should be distributed among administrators and staff. Leadership in successful programs is not one sided; it is both top-down and bottom-up (Fullan, 1993a). Fullan adds that empowering teachers as leaders broadens the leadership capacity in a system and expands the principal's leadership role.

The Community and Culture of the Schools Are Taken Into Account

Successful policy makers respect and address the attitudes, needs, and self-interests of people affected by change and understand the climate and culture into which changes are to be introduced. If the community is to be committed to, and not simply cooperated with, proposed changes, time and patience are required of leaders in the change process so that they can understand and respond to resistance (Sarason, 1990).

There is evidence that it helps to involve parents and keep the community informed of changing practices. School leaders also need to include staff members who are not directly involved in a new program. Ignoring non-users in the school community will lead to resentment (Fullan, 1985).

Leaders Understand That Change Takes Time and Is Often Met With Resistance

Change is a process, not an event, and as such takes time (Fullan, 1985). Because it is developmental, change requires practice and feedback. The first year in which a change is introduced may be particularly difficult and the introduction of change creates more turbulence (Fullan, 1985; Huberman, 1983). It is crucial that participants in, as well as observers of, a new project, understand the change process and are aware that problems and discouragement are inevitable in the initial phases of the program. Support for teachers from their peers and leaders is essential at this time. Patience is indispensable.

Conclusions

To ensure their lasting success, it is necessary to reduce the costs of changes and maximize the benefits for those who must subscribe to the change. The process should not be entered naively, but realistically, with the understanding that it may be prolonged and difficult. Because change is difficult, the needs and interests of the people involved must be taken seriously. Teachers and other personnel affected by change must be provided with ample support in terms of time, resources, collaboration, and leadership. Everyone affected by the change, students, parents, non-participant staff and students, and the local community, should be considered. Leaders and policy makers must have a thorough understanding of the change process and be guided by clear principles, but they must also be open and flexible enough to build a vision that is shared and practical.

Benefits of a Multicultural Approach to the Education of Minority and Low Socioeconomic Status Students

Definitions

Multiculturalism may mean different things to different people. Five approaches to multiculturalism are commonly depicted in the literature: education of the culturally different, ethnic studies, human relations, multicultural education, and education that is multicultural and social reconstructionalist (Sleeter & Grant, 1994).

The goals of educating the culturally different are defined as increasing minority access to the mainstream through the use of language, curriculum, and materials that are culturally relevant to the target students. This may involve the inclusion of modes of instruction that account for the varied cognitive learning styles of different ethnic and cultural groups (Banks, 1993; Shade, 1994; Sleeter & Grant, 1987).

The purpose of ethnic studies is to teach about ethnic groups. Courses are usually added to the curriculum or substituted for other courses. This approach is not commonly advocated by multiculturalists because it lacks depth and does not result in meaningful change in the curriculum (Grant & Sleeter, 1985).

Multicultural approaches that focus on human relations attempt to increase sensitivity and promote tolerance between different ethnic groups, as well as develop a positive self-concept. Grant and Sleeter (1985) use the term "multicultural education" to describe what they consider to be the most common approach to multiculturalism. Advocates of this approach believe that teachers "should help students develop ethnic self identities, knowledge about different cultural groups, respect for others' right to be different, and competence in more than one cultural system" (Grant & Sleeter, 1985, p. 101).

Education that is multicultural and social reconstructionist focuses more on the relationships of the oppressed and the oppressors and seeks to actively challenge unequal relationships. Students are taught to analyze oppression and take a role in the change process (Grant & Sleeter, 1985).

Confusion can arise when people use "multiculturalism" differently, possibly referring to different educational approaches without realizing it. It helps to have an overarching definition of multiculturalism that encompasses the various definitions, one or more of which may be incorporated into the multicultural programs of a school system (Grant & Sleeter, 1985). A multicultural outlook is manifest in any program that seeks to empower students through recognizing cultural diversity as positive and as a strength.

Curriculum Content

Addressing the area of curriculum content, there are two curricular approaches that target the changing world around us: global education and multicultural education. They both seek to: (a) improve interpersonal and intergroup relations; (b) increase awareness of the impact of global and national forces, trends, and institutions on different groupings of people, including national and ethnic groups; (c) reduce stereotyping and increase intergroup understandings; (d) help students comprehend the significance of human diversity, while at the same time recognizing underlying, globe-girding commonalties; and (e) improve intercultural communication (Zuck, 1983). Global education serves to introduce students to part of a world mosaic, while multicultural education seeks to explore current volatile social and political issues among cultures represented in this country. Although a study of a culture from the distant past may be less threatening, it is neither contemporary nor relevant (Wong, 1993).

Multicultural Education as Early Intervention or Prevention

Elementary teachers are more inclined to work with multicultural curriculum than secondary teachers because the former tend to lack the degree of concern about content coverage that often typifies the latter (Sleeter & Grant, 1987). Moreover, issues inherent in a multiculural curriculum, issues of culture and race, should be introduced in a child's early formative years. Research suggests that children's notions of prejudice, bias, and racism are learned at an early age and early interventions may counter these (Derman-Sparks, 1989).

The cultural make-up of the classroom has undergone explosive changes and researchers continue to examine the incompatibility between culture and the school classroom (Larkin, 1993; Perry & Fraser, 1993). Therefore, there is a need for increased understanding and respect for children's cultural backgrounds and home influences in the school environment with curricular accommodation in content, modes of teaching, teacher/student verbal interactions, students' preferred learning styles, and classroom climate (Banks & Banks, 1993; Guild, 1993; Tharp, 1989).

Cognitive Styles, Ethnicity, and Socioeconomic Status

Research on learning styles indicates that there are a variety of ways different students approach learning tasks. Banks (1991) asserts that certain learning and motivational styles tend to prevail among ethnic and/or socioeconomic groups. While there is much diversity within ethnic and socioeconomic classes, class and ethnicity interact to influence learning styles. It is important to avoid stereotypes and generalizations about individuals within cultural groups. Nevertheless, educators need to understand that some minority groups tend to approach learning tasks in ways that conflict with the rules and expectations of the traditional academic setting (Banks, 1991).

According to Banks (1988) cultural adaptations of African Americans may have resulted in learning styles that are different than those most conducive to achieving in current classroom settings. Conflicts between learning, motivational, and cognitive styles of some ethnic minorities and the styles fostered by the schools can result in lower achievement among these minority students (Banks, 1988; Shade, 1989).

Field independent and field sensitive cognitive styles can significantly affect student success (Banks, 1988; Shade, 1981). These terms refer to methods by which people approach perceptual tasks. Field independent people are good at cognitive restructuring and information transformation and, as a result, rely more on their own perceptions rather than on external referents. Field independent learners are also more comfortable with abstract analysis and impersonal learning situations. They are generally more object-oriented and work well in isolation.

Field sensitive (also referred to as field dependent) learners have difficulty distinguishing necessary parts to solve a problem, refer to external cues to interpret information, and tend to be more concrete thinkers. Integrative and interactive cooperative learning programs are more appropriate to field sensitive learners. They are generally more person-oriented and learn through inter-personal interactions. Researchers report that African American, Hispanic, and other minority groups are more likely to be field sensitive than field independent learners. Children of low socioeconomic status are also more likely to display field sensitive learning characteristics (Banks, 1988; Shade, 1981).

Cognitive skills involving analysis and classification have been shown to be influenced by ethnicity and socioeconomic status. Minority and low income students have been found to classify objects by their relationships to each other, whereas middle class Caucasian students generally classify objects based on similarities in specific elements (Shade, 1981).

Regardless of ethnicity, students of low income background tend to be more impulsive and exhibit an external locus of control (Banks, 1988; Shade, 1989). Research by Gilbert and Gay (1989) indicates that low socioeconomic status students are less able to see a relationship between their behavior and consequences. Motivation and confidence are also affected by locus of control. Teachers often expect students to display only cognitive engagement in a learning task. In contrast, Gilbert and Gay found African America students were usually physically and affectively engaged in learning activities. Conflicts between teacher expectations and student behavior resulted in unfortunate misunderstandings and tensions between teacher and students. The end result of these conflicts may be energy wasted on non-academic tasks.

Banks (1993) and Shade (1981) have found that minorities and low socioeconomic status students function best in cooperative, interactive, and loosely structured environments where teachers and students work together toward a common goal. If learning experiences are to have a positive outcome for these students, there needs to be a focus on interrelationships. "Black students perform better when immersed in a system of interrelated learning encounters; the more that content, context, procedures, and climate reinforce one another, the more thorough and enduring their learning will be" (Gilbert & Gay, 1989, p. 281).

It is also noteworthy that spoken, not written, language is the primary mode of communication among African American students (Gilbert & Gay, 1989). African American students tend to be more dramatic than direct in their approaches. Responsive teachers will set high standards for African American students, but also consider the differences in styles and provide African American students with opportunities to translate from the spoken to the written and from indirect to direct modalities of learning. Educators must build on students' primary learning modes as well as help learners incorporate different modes into their styles (Gilbert & Gay, 1989; Grant & Sleeter, 1985).

Providing All Students a Relevant Context for Learning

Regardless of learning styles, children from minority cultural backgrounds may become disengaged from learning if they fail to see the relevance of the curriculum to their own experiences. Incorporation of a student's language and culture into the school community improves the quality of interactions between the student and the school community (Cummins, 1986).

An initial step in transforming the curriculum into a multicultural perspective is carried out through the use of multicultural books (Banks & Banks, 1993). The sharing of diverse pieces of literature, such as an author's view of life and world events, also lends itself to an inclusive multicultural approach (Willis, 1994). Since more traditional curricular offerings fail to mirror the nation's present diversity, these additions present a perspective to better prepare students for a future in a multicultural world (Vann & Kunjufu, 1993). Such use of historical and fictional literature from cultures in a classroom enables students to identify themselves and their races in their studies. The appeal of learning materials can be expected to improve a child's motivation to become engaged in activities (Gilbert & Gay, 1988; Walker-Dalhouse, 1993).

Integrative and culturally relevant reading programs are needed to motivate students and to incorporate modes compatible with field sensitive learning styles (Gilbert & Gay, 1988; Walker-Dalhouse, 1993). "If the interest and appeal of the materials used to teach reading has a positive effect on learning, and if African American youngsters find the content of the basal readers uninteresting and meaningless, then it is senseless for schools in Black communities to waste scarce economic resources to purchase these books" (Gilbert & Gay, 1988, p. 281).

Making a connection between a child's home and school culture is not simply a matter of inclusive literature. Literally incorporating a child's culture into schooling by encouraging parents and the local community to participate actively in the school is important. Minority parents are interested in promoting academic performance but often feel excluded from the school community (Cummins, 1986).

Students from non-dominant cultures will be empowered in the school context to the extent that the communities themselves are empowered through interactions with the school. When educators involve minority parents as partners in their children's education, parents appear to develop a sense of efficacy that communicates itself to children, with positive academic consequences (Cummins, 1986).

Community involvement may take myriad forms. In the Washington State Yakima Equity Study (Vasady & Maddox, 1992), parent involvement included helping with homework, joining the Parent Teacher Association and Parent Advisory Councils, volunteering in classrooms, attending parent nights, going to class and school parties, talking with classes about their cultural heritage, serving as parent liaisons, and attending athletic and cultural performances at the school. Staff actively promoted parent and community involvement by holding open houses and hosting other activities by accommodating language and literacy differences among parents, by making home visits to families who did not have easy access to school functions, by regularly sharing information about programs, problems, and progress, and by encouraging parents to take a leadership role in the school community. Even if a child's own parents are unable to become actively involved in the school, she or he will benefit from the inclusion of adults with a similar background in the school setting (Vasady & Maddox, 1992).

Multiculturalism and Self-Esteem

Providing a relevant context allows for personal connection between the child and the learning environment, thus creating a more meaningful experience for the learner. The result of using meaningful content for personal connections is especially important for minority children for reasons of self-esteem.

Self-esteem has historically been associated with educational achievement (Ford, 1994a). Self-esteem is the evaluation an individual makes of his/her self-concept (Verma & Mallick, 1988). Development of a positive or negative academic self-concept often occurs as soon as a child enters the school environment.

> When a student with low self-esteem enters a classroom, self-concept becomes one of the most challenging individual differences in how he or she will learn. School becomes an arena for failure that prevents them [sic] from achieving the success needed for high self-esteem. (Bennett, 1986, p. 131)

Researchers differ on whether teaching with a multicultural perspective either builds self-esteem or inhibits self-esteem from eroding. However, instruction with curriculum that validates a respect for all cultures communicates a model especially useful for young children's development of self-esteem by developing contact with each other while retaining their ethnic identity (Frisby & Tucker, 1993). Students' perceptions of self-esteem are critical, but they are especially crucial for minority students. "The culturally diverse learner who feels that others perceive him or her as inferior, deficient in some manner, or in stereotypical terms, may begin to feel less than adequate, or unable to cope with the United States school system" (Baruth & Manning, 1992, p. 231).

Research specific to minority populations and self-esteem is important to the current study. An area of research on Afrocentric curriculum materials maintains that, especially for African American children, positive images of African American people and world events enhances self-esteem (Murell, 1993; Vann & Kunjufu, 1993).

Outcomes of Multicultural Education

While the literature endorsing multicultural education is vast, there has been little empirical research into the effectiveness of multicultural approaches in meeting goals. Perhaps this is because the goals are broad and many of them do not lend themselves to quantitative measurement. Goals of justice, relevance, and inclusion are hard to break down into measurable objectives; however, it should be possible to demonstrate that components of multicultural education have had an impact on student outcomes.

Some studies have demonstrated gains in student achievement as a result of applying multicultural approaches to education. Some of these approaches have integrated many aspects of multiculturalism, and therefore it is difficult to make inferences as to what aspects may have been particularly effective. The Arkansas Multicultural Reading and Thinking (MCRAT) (Arkansas Department of Education, 1991) program reported significant gains in analytical thinking among participating students. The multicultural emphasis of the program was intended to provide a meaningful context for learning through literature and parent and community involvement. Because other treatments included direct instruction in higher-order thinking skills and no control or cross group studies were done, it is not possible to ascertain what impact the multicultural treatment alone had on the results.

A Hawaiian reading program changed instructional methods to meet the more collaborative learning styles of native Hawaiians. An interactive approach to basal reading stories was designed to bring the stories into a frame of reference the children could understand, and to approach the story in a manner consistent with Hawaiian storytelling patterns. Reading and verbal intellectual abilities were reported to have improved dramatically (Au-pei-Hu, 1980).

The Haringey project (Tizard, Schofield, & Howison, 1982) in a multicultural district focused upon a collaborative relationship between school and the home. In this controlled study, parents listened to children read at home and interacted with school staff often. Comparison groups included students who read at school to expert reading teachers and students that received no treatment. Highly significant progress was made by children who read to their parents, regardless of whether those parents spoke English or not, as compared to children in other groups. In addition to significant improvements in performance, children in the home collaboration group were reported to have demonstrated increased interest in school and improved behavior (Tizard, Schofield, & Howison, 1982).

Conclusions

Prevalent school structures and practices place minority and low socioeconomic status students at a disadvantage in the classroom. A multicultural program that empowers diverse learners is needed to allow each child to achieve at his or her full potential. Such a program must recognize and provide instruction suited to varied learning styles. Students will be drawn into the school culture as they see themselves and their heritage reflected in their studies and in their daily encounters in the school building.

Mentoring

Mentoring is a widely recommended strategy for meeting the intellectual, social, and emotional needs of gifted students (Cox, Daniel, & Boston, 1985; Frey, 1991; Gallagher, 1985; Hollinger, 1991; Torrance, 1984). The practice of pairing an older, more experienced person with someone who is younger or less experienced dates back to ancient times (Anderson & Shannon, 1988). Today, mentoring has gained popularity in a variety of professional and academic settings with various student and adult populations. This section of the review examines the literature relating to gifted and talented individuals, focusing on programs designed for students in grades K-12.

Programs With an Academic Focus

Mentoring programs provide opportunities for gifted students to work in one-on-one relationships with experts to provide learning experiences of a kind and depth not ordinarily available in the traditional classroom (Beck, 1989; Hendricks & Scott, 1987;

Lupkowski, Assouline, & Vestal, 1992; Milman & Schwartz, 1992). Gifted programming models that emphasize the development of independent student projects, such as the Enrichment Triad/Revolving Door (Renzulli, 1977; Renzulli, Reis, & Smith 1981), readily encourage the use of experienced individuals to serve as guides to young investigators as they pursue varied interests. Fast-paced programs designed to accommodate highly able math students at their own rate of learning, such as the Study of Mathematically Precocious Youth developed by Julian Stanley at Johns Hopkins University, and the Investigation of Mathematically Advanced Elementary Students developed at the Connie Belin National Center for Gifted Education, rely on mentors to individualize programs for highly able students (Lupkowski, Assouline, & Vestal, 1992).

In addition to providing instructional opportunities in specific content areas (Clifford, Runions, & Smyth, 1988; Minnehan & Strunk, 1992; Pizzini, 1985), mentoring relationships can permit students to explore a wide variety of interests related to potential careers (Comer, 1989). For example, mentoring programs for young children can focus on the learning characteristics of young children. One such program developed for young learners was designed to encourage and foster children's inquisitiveness and heighten their curiosity through mentoring (Hendricks & Scott, 1987).

Programs With a Social Development Focus

Although mentoring programs for the gifted typically focus on providing for students' cognitive needs, the need for role models who can guide the development of decision-making skills and the development of self-esteem also is recognized (Frey, 1991). Access to community leaders and professionals serving as mentors can provide gifted students with opportunities to develop leadership and interpersonal skills (Milman & Schwartz, 1992).

Other programs focusing on social development have been designed to provide consistent caring relationships for students who are at risk of failing or dropping out of school (Blum & Jones, 1993; Freedman, 1989; McPartland & Nettles, 1991; Slicker & Palmer, 1993) and, therefore, are not specifically designed for gifted students.

One program that specifically targets at-risk gifted students is Project Synergy (Borland & Wright, 1994) conducted through Teachers College at Columbia University. Although programs usually provide adult mentors for students, Project Synergy invites gifted adolescents to mentor younger, potentially gifted students (Wright & Borland, 1992). As mentors, adolescents provide coaching in basic behaviors associated with success in school such as learning to raise a hand to be called on by the teacher. In this way, the program facilitates the learning of skills often possessed by more advantaged students.

Benefits of Mentoring

Much of the literature on mentoring provides descriptions of programs. Benefits reported by those who have been mentored include direction in career planning, an increase in content knowledge and interpersonal skills, development of talents and personal standards, and enhancement of self-esteem, self-confidence, and creativity (Edlind & Haensly, 1985). However, little empirical evidence relating to the effects of mentoring exists (Edlind & Haensly, 1985; Merriam, 1983). Data that are collected often consist of informal evaluations and after-the-fact testimonials in which successful people reflect on the advantages of having a mentor (Collins & Scott, 1978; Merriam, 1983; Roche, 1979). This may be due to the fact that the impact of mentoring may be detected years after the process began. For example, in his 22-year study, Torrance (1984) found a significant relationship between (a) having a mentor and completing school and (b) having a mentor and high creative achievement. Kaufmann, Harrel, Milam, Woolverton, and Miller (1986), in a survey of 1964-1968 Presidential Scholars, found that gifted individuals rated role modeling, support, and encouragement as the most important functions mentors served.

Mentorships as a Factor in Talent Development

The retrospective research on talent development suggests that mentoring provides students with both the academic and social benefits critical to talent development (Arnold & Subotink, 1995; Berger, 1990; Bloom, 1985; Csikszentmahalyi, Rathunde, & Whalen, 1993; Gardner, 1993b; Goertzel, Goertzel, & Goertzel, 1978; Kerr, 1985; Piirto, 1994; Torrance, 1984).

Mentoring programs have also been studied by collecting immediate feedback from participants in programs on various aspects of the experience. Survey data collected immediately after completion of a program can provide insight into the short-term effects and perceived strengths of that program. For example, at the mentoring program offered by the Gifted Education Research Institute at Purdue University for students in grades 4-12, 93% of the students were rated as having worked well independently ($n = 44$) and 75% of the students wanted to participate in the program again (Ellingson, Haeger, & Feldhusen, 1986). The majority of students also reported being motivated by their mentors to study an area further.

The findings of Project Synergy (Borland & Wright, 1994) illustrate the difficulty in evaluating the mentoring component in isolation. The first conclusion in the evaluation of the project points to the importance of considering the long-term impact of mentoring. It is likely to take several years before one can assess the academic success of mentored students in terms of such factors as grades, persistence, and aspirations. Second, mentoring is usually one component of a multifaceted program. The direct effects of mentoring or any one component are often difficult to isolate.

Guiding Principles for Project START

The literature reviewed above was used as the basis for several aspects of Project START. The background provided on the problems of identifying non-traditional gifted students led us to use a performance assessment strategy for identification. The importance of early identification and nurturance of talent led to a focus on grades K-2. Finally, the literature on constructivism, multiculturalism, and mentorship was used as the basis for curricular modifications and the inclusion of a mentorship component.

CHAPTER 2: An Overview of Identification and Placement of Students and Interventions

The research described in this report emanated from a collaborative relationship between the Charlotte-Mecklenberg Schools (CMS) in Charlotte, North Carolina and the University of Virginia site of The National Research Center on the Gifted and Talented (NRC/GT). Both the conception and implementation of the project were collaborative. While the NRC/GT provided consultation and guidance in sources for the development of the identification protocol and provided the vast majority of the staff development program, the overall administration of the project was directed by the staff at CMS.

Sixteen schools ($n = 16$) participated in the project based on agreements by the principals to participate in the project as part of their school improvement plans. The schools themselves were initially classified according to whether they were perceived as having a climate that could be characterized as higher school achievement motivation or lower school achievement motivation. This distinction was made on the basis of examining questionnaires that were collected from parents of the CMS school division and central office personnel, and a ranking by faculty who supervised student teachers in the schools in the study. A split of schools into the "lower" and "higher" groups was indicative of relative rating, not absolute qualities. Initially, one-half of identified experimental (treatment 1 and treatment 2) and control group students were selected from eligible participants in "high" motivation schools and one-half from eligible participants in the "low" motivation schools. Extremely high turnover of administrative staff during the first two years of the project, and consequent changes in school climate made this classification a meaningless variable in the study. Details of the rating process and classification of the schools are available on request.

The staff at CMS and the NRC/GT determined that the project would focus on only three (verbal-linguistic, logical-mathematical, spatial) of the seven intelligences identified by Gardner (1983). Data on interpersonal intelligence were collected for construct validation purposes only. The interpersonal dimension was not targeted in the intervention. The process of carrying out the selection of students and the assignment to experimental and control conditions are explained in the section that follows.

The two institutions also determined that the focus of the curricular intervention efforts would be (a) the use of multicultural curriculum, (b) active learning with a heavy emphasis on manipulative learning as a basis for introduction to concepts, and, ultimately, abstract thinking using those concepts, (c) language development to prepare students for identification for the gifted program at grade 3, and (d) the conceptions of intelligences offered by Howard Gardner (1983) as the basis for extending the types of learning tasks generated by teachers for use in their classrooms. Further, because of the literature suggesting the importance of involving families in the educational programs of their children, a family outreach program was included as part of the intervention for all identified students in the experimental groups. A mentor program was provided for half of the students in experimental conditions. The groups were then considered Experimental 1 and Experimental 2.

Assessment, Training, and Student Selection and Placement

During year one, one aspect of the staff development effort and training focused on the creation of a protocol of assessment tools to be used in identifying students deemed to be exceptionally able along any one or more of four dimensions described by Gardner (verbal-linguistic, logical-mathematical, spatial, or interpersonal). This protocol contained a series of performance tasks including solving tangram puzzles, the disassembling and re-assembling of a drain, the creation of oral stories using props, and the calculation of the numbers of persons getting on and off a bus on a simulated trip. Teacher ratings of aptitude along each dimension were also included in the protocol. The tasks and teacher ratings were adopted or adapted from materials that had been developed by the staff of Project Zero at Harvard, by the Montgomery County Schools (Montgomery, Maryland) or by C. June Maker as part of a model project she was conducting also under the auspices of the Javits program. The battery was developed during the course of more than 30 hours of training provided to 16 first-grade lead teachers, 3 kindergarten teachers, and 4 support team members with instruction and leadership provided by Carol Tomlinson (START consultant and NRC/GT project co-director), C. June Maker and Judy Rogers from the University of Arizona, and Mara Krecevsky and Julie Viens from Montgomery County, MD.

A cadre of school personnel including teachers, the University of Virginia project director, and other school staff went to each kindergarten and first-grade classroom to administer the battery of instruments to the students in the 16 schools selected to participate in the project. Project staff reviewed the results of the administration of the battery combined with teacher ratings based on the specially designed checklists and identified students who met the criteria established by the staff for selection as a START student. Students having talents in one or more of the following areas were identified as START children: verbal-linguistic, logical-mathematical, or spatial. Although children were assessed on the personal intelligences, they were not identified or placed as START children on the basis of those assessment indicators. A subset of 256 identified children were randomly selected and placed in clusters of eight in the classrooms of 32 first- and second-grade teachers who had been selected as the "START teachers" in each building (one first-grade and one second-grade teacher in each of 16 schools). Eighty-six ($n = 86$) children were randomly selected as controls and placed in classrooms of "non-START" teachers in the same schools. A third group of children who were very close to meeting the criteria were also placed in both START and non-START classrooms and identified as START children to the teachers. They were identified as foils in the original design, but attrition prohibited meaningful analysis using this group of children (see Figure 1).

The original proposal did not include plans for an evaluation of the assessment protocol; however, the progress of the project, as well as concerns about interpretation of the student outcomes, seemed to be predicated on an understanding of the psychometric qualities of the instruments used to select the students.

In one half of the schools the students received the intervention treatment which included curriculum modifications and family outreach. In the other half of the schools, the children also were provided with mentors along with the curriculum modifications and family outreach.

Nature of the Intervention

Specific Training of START Teachers

START teachers participated in five types of project-related sessions. During the first year, first- and second-grade teachers took part in two- to three-hour orientation sessions designed to help them gain comfort with the broad goals of START, its various components and players, and their own roles in it. Second, they were invited and encouraged, but not required, to attend sessions in their schools conducted throughout the project period for START families by the family outreach coordinator. Third, START teachers came to monthly staff development and information sessions conducted on a districtwide basis by the START coordinator. These sessions were generally 1 1/2 hours in length and included dissemination of information about the project (e.g., budget, ordering of materials, discussions of upcoming special events for START staff and/or families), brief presentations and discussions on the instructional goals and elements of a START classroom, and interactions among attendees to share teaching ideas and concerns. Specific staff development in the sessions was typically led by START consultant Dr. Carol A. Tomlinson. A fourth type of staff development session was provided for START teachers during one or two full-day, districtwide sessions. START teachers were given released time from their classrooms for these START workshops. These sessions allowed extended time for study and discussion of project instructional elements, interaction with local experts and resource people related to START goals, and hands-on, collaborative design of START curriculum and materials. These same opportunities were provided during the second and third years for third-grade teachers.

The fifth type of staff development took place during START summer institutes. In year one of the study, the institute took place for approximately six hours daily over the course of five days. Participants (first- and second-grade START teachers) began a study of goals of START, elements of a START classroom, and community resources which could assist in moving toward project goals. During the second summer of the study, first- and second-grade teachers worked daily for five days on developing concept-based curricula for their classrooms. Also during the second summer institute, third-grade teachers, all newly recruited for the year ahead, took part in sessions designed to orient them to project goals, previous learning experiences of START students with whom they would begin working, and ways in which progress toward START goals could be continued in their classrooms. In both of the summer institutes, the START consultant provided lectures, demonstrations, and guided practice in instruction in a START classroom. Other staff development providers during the summer institutes included national and local experts on particular elements essential to START (e.g., establishing a multicultural environment), local educational leaders (e.g., Director of Programs for the Gifted), and local teachers selected for their ability to talk about and model key elements of rich instruction (e.g., whole language, use of manipulatives in the primary classroom). A third summer institute focused on drafting concept-based curriculum units within a multiple intelligences framework.

Elements of START Instruction

Four key elements of instruction were described in the project proposal as pivotal to instruction in a START classroom.

The first element was based on principles that guide good constructivist teaching including:

- students should work with questions and problems relevant to them;
- learning should be based on primary concepts of the topic/discipline being explored;
- students should be assessed in the context of learning; and
- students should make meaning of key ideas and come to own them (or borrow them) for their own.

During Project START training, these principles were at the core of direct instruction of the teachers, modeling for the teachers, observations and coaching, and curriculum development. Thus, we worked with teachers to (a) organize lessons that would be relevant to START learners, (b) create lessons that would help students develop a sense of organization and meaning of subjects explored (vs. coverage of discrete data), (c) develop patterns of classroom interaction that encouraged students to discover key principles that govern areas of study, (d) conduct "generous" classroom discussions that put students on center stage and that supported students in discussing their problems and questions related to key ideas being explored, (e) use manipulative materials in all subject areas so that students had to sort out and apply meanings, (f) use materials and talks that felt familiar to students from culturally diverse and/or low SES settings, and (g) be consistently observant and reflective about student work so that student effort provided teacher directions about next steps in teaching individuals or small groups of individuals.

Using this framework, the teachers were instructed in how to use a performance orientation to instruction or student use of manipulatives as a vehicle for understanding (that is, students actively constructing meaning based on constructivist principles). The intent of incorporating this approach into the classroom instruction was to allow students to work from concrete learning experiences to more abstract ones, to allow students to use the manipulation of materials so as to enable them to construct their own meaning and patterns of meaning. Consistent with constructivist theory, hands-on learning promotes a sense of the usefulness and applicability of ideas and information. The teachers were instructed to use activities in which students applied learning, used learning organizers, were engaged in role-playing, drama, and the solution of real-life problems in their own schools, homes, or communities. Teachers were also provided strategies for introducing students to how various professions and occupations used ideas and skills in the curriculum.

The second area in which teachers were instructed to modify curricular offerings in their classrooms was in the use of formal language orientation. The intent was to immerse the children in the use of standard English to prepare them for later success in the dominant language. The teachers were provided strategies, using a whole language approach, to tap student interest, build on student experiences, extend student exposure to ideas and information, and provoke fluent thought at high levels, which both valued student expression and were likely to foster school success. Strategies such as extensive and on-going conversations between and among students as well as students and teachers, pre-writing and writing activities, interviews, question production, oral reading and silent reading, storytelling, and inquiry learning were included as part of the staff development. The premise was that students should be consistently at the center of language use, engaging in activities such as storytelling, explaining reasons, sharing experiences, planning aloud, and comparing and contrasting.

A multicultural orientation to curricular development—including use of familiar materials, focus on relevant experiences, concrete learning opportunities, consistent exposure to multi-ethnic materials and people, an active learning environment, establishing peer support and cooperation, building appreciation for varied views of events—was designed to assist students in realizing that all cultural and ethnic groups have contributed significantly and continue to contribute significantly to human knowledge and experiences. The teachers were encouraged to create experiences wherein students were able to identify key ideas and concepts across cultures, disciplines, and times, as well as learn about contributors to knowledge from varied ethnic groups. Teachers were provided resources that included materials reflecting diverse cultures, along with ideas for concept-based instruction that helped them make connections across fields of study and culture. Role models also were provided.

The instructional intervention component of START guided teachers in continual development of understanding and application of these elements, within the context of effective teaching in general. Teachers and staff developers worked with the image of a classroom in which teachers begin with a sense of what is important for students to learn in the way of information, concepts, principles, and skills (content)—as well as ways in which students might come to "own" the ideas (process) and demonstrate their proficiency (products). From that beginning point, effective teachers would "filter" and modify their sense of teaching by examining their learners and teaching context. For example, factors such as student attention span, student readiness, daily schedule, and accessibility of materials help a teacher take her original sense of subject matter and sculpt it in ways most likely to give learners access to and comfort with that which should be learned.

START teachers were asked to focus their instructional plans by using one additional planning "filter." As the "general" filter of expectations was employed to help match instruction to general learner need, teachers were assisted in considering particular profiles and needs of START learners, and in using the "START filter" to guide their use of the four START elements in ways which might address specific START learner needs (see Appendix A for a schematic model). The assumption was that teaching a START lesson did not supplant, but would rather augment or extend effective teaching.

Family Outreach

During the first summer institute in which teacher training was provided there was also specific training for family outreach coordinators in each school on the characteristics of culturally diverse and economically disadvantaged gifted students, strategies for training parents to recognize, nurture, and maximize their children's talents, and ways to involve parents in the education process. During the second year of the project, there were monthly 11/2 hour meetings of the 16 family outreach coordinators with the project director at CMS. In the third year of the project, a one day family empowerment seminar was attended by family outreach coordinators and community representatives during the summer training program and bi-monthly meetings to discuss strategies and plan activities. Each family outreach coordinator developed her own plan for community involvement, although some activities, such as a museum exploration day for parents and their children, was a joint venture across all schools.

Mentor Development

A mentorship component was developed in each of eight schools. Community members volunteered to visit classrooms of START students and work with them as tutors, advisors, or simply "friends who cared and encouraged." Mentors visited START students three or four times a year to once a month for the duration of the project.

CHAPTER 3: Methodology and Procedures

The research project was divided into several components. One component focused on an investigation of the psychometric properties of the assessments used to select children for the START project. These assessment tools were adaptations of instruments developed by the researchers associated with Howard Gardner and others who have worked on multiple intelligence assessments. The second aspect of the investigation focused on the quantitative changes in students involved in the project. These quantitative outcomes were assessed in terms of both test score change and the increased likelihood of identification as gifted. The qualitative component of the project was also divided into two components, the first being an in-depth look at students targeted as highly successful or unsuccessful in the project and the second focusing on teacher change.

Research Questions

Specific research questions focused on the evaluation of the psychometric properties of the instruments used in the identification process included:

1. Do the MI-based assessments show evidence of reliability and validity?
2. Are there ethnic and/or gender differences in student performance on the MI-based assessments?
3. How do the MI-based, alternative assessments compare to traditional, standardized assessments with respect to psychometric properties?

Psychometric Properties of Alternative Assessment Battery: Cohort 1

Sample

The sample for this portion of the study consisted of 1,813 children enrolled in kindergarten and first-grade in 16 Project START schools at Charlotte-Mecklenburg Schools (CMS) in Charlotte, North Carolina during the 1992-1993 school year. The target population for assessing the instruments was ethnically diverse and/or low socioeconomic status (SES) kindergarten and first-grade students. Female students comprised 48.2% (n = 873) of the sample, and ethnic composition of the students was as follows: Caucasian, 18.8%; African American, 71.3%; Asian American, 1.8%; Hispanic American, 2.5%; and other ethnic groups, 3.5%. SES was determined by participation in the federal government's free/reduced lunch program. The lunch status could not be determined for 59.8% of the sample because of issues of confidentiality of data for those not selected for the project. Of the remaining 729 students, 48.4% (n = 353) received free or reduced lunch.

Validity Sample

Data on additional measures of attitude, self-concept, and achievement were obtained for the 371 children who were selected for the talent development program and were randomly assigned to experimental or control conditions. In the following sections that report evidence of concurrent validity, only results from students in the validity sample were used. As these students generally received higher scores on the assessment activities, the results of the analyses which employed the restricted sample should not be generalized to the entire sample. In the validity sample, 49.6% (*n* = 184) of the students were female, and the ethnic composition of the students was as follows: Caucasian, 23.5%; African

American, 63.6%; Asian American, 6.5%; Hispanic American, 2.7%; and other ethnic groups, 3.8%. Nearly half of the students (48.2%, *n* = 179) received free or reduced lunch.

Instrumentation

In the spring of 1993, students were assessed using the Multiple Intelligences Assessment Technique (Udall & Passe, 1993), which is based upon the work of Project Spectrum at Harvard (Gardner & Hatch, 1990) and that of C. June Maker at the University of Arizona (Maker, 1993; Maker, Nielson, & Rogers, 1994; Maker, Rogers, & Nielson, 1992) with local modifications. The battery of assessments consisted of 13 performance-based activities, teacher ratings, and observational checklists corresponding to four of the multiple intelligences: logical-mathematical, linguistic, spatial, and interpersonal. The 13 checklists, ratings, and activities are summarized by subscale in Table 1.

For each activity, student performance was rated on a scale as "not evident or not observed" in a given setting (0), "evident" (a), or "extremely evident" (b). Teachers and external observers received training on administration of the assessments, and guidelines were provided with respect to typical behaviors that should be rated 0, 1, or 2 for each assessment activity (Udall & Passe, 1993). In order to obtain estimates of construct (discriminant and convergent) validity, particular Iowa Tests of Basic Skills subtests (ITBS; language arts, mathematics, reading comprehension, and vocabulary) were administered.

Table 1

Activities Used in the Alternative Assessments Categorized by Intelligence for Cohort 1

Activity Title	Type of Assessment
Spatial intelligence	
1. spatial checklist	teacher rating
2. Pablo	students construct 3-D animal from puzzle pieces
3. mechanical pump	students take pump apart, put it back together
4. tangrams	students manipulate puzzle pieces
5. artwork	students draw or paint a picture
Logical-mathematical intelligence	
6. math-logical checklist	teacher rating
7. bus activity	board game
8. math activity	students solve mathematical problems
Linguistic intelligence	
9. linguistic checklist	teacher rating
10. storytelling activity	students tell a story using various objects
11. pictorial writing prompt	students write or draw a story
Interpersonal intelligence	
12. interpersonal checklist	teacher rating
13. interpersonal skills	observation checklist

Data Analysis

Reliability

Because of logistical problems associated with large-scale performance assessments, test-retest and inter-rater reliability could not be obtained for the performance assessments, although obtaining this evidence is a priority of future administrations of the assessments. Cronbach's alpha was calculated as a measure of internal consistency for each of the four subscales (verbal-linguistic, logical-mathematical, spatial, and personal).

Validity

Correlations among the assessment subscales and ITBS subscale scores were computed to obtain evidence of construct validity. Confirmatory factor analysis was used to determine whether the activities assessed the corresponding four intelligences (verbal-linguistic, logical-mathematical, spatial, and personal). An examination of convergent and discriminant validity was conducted using multitrait-multimethod matrices.

Gender, Ethnic, and School Differences

Fixed effect analyses of variance (ANOVAs) were conducted using the ITBS subscale scores as dependent variables to determine the presence of any differences based on gender, ethnicity, or school. Eta squared (η^2), the percent of variance accounted for by significant effects, was calculated as a measure of effect size (Tabachnik & Fidell, 1989), and recommended guidelines for interpreting effect sizes (Rosenthal & Rubin, 1979; Rosnow & Rosenthal, 1988) were followed.

Psychometric Properties of Alternative Assessment Battery: Cohort 2

Sample

The students involved in cohort 2 (n = 1,077) attended kindergarten during the 1993-1994 school year. The sample was roughly balanced by gender (51.1% female), with the following ethnic composition: Caucasian, 27.0%; African American, 66.6%; Asian American, 1.9%; and Hispanic, 4.5%. A majority of the students were considered to be economically disadvantaged by school district personnel.

Instrumentation

The assessment battery used with cohort 1 was modified for use with cohort 2 during spring of the 1993-1994 school year (Table 2). Two observer scales were added to provide ratings of each student's use of linguistic skills during the assessment process and use of interpersonal skills during other classes such as music and physical education. One performance activity (disassembly and reassembly of a mechanical pump) was replaced with another (disassembly and reassembly of a mechanical drain). These changes were made to incorporate an opportunity for more data from more sources (addition of observation scales) and to address teacher complaints about the difficulty of the pump activity. With the modifications, the battery included four classroom teacher rating scales (activities 1, 6, 9, and 13), three observer or non-classroom teacher rating scales (12, 14, and 15), and eight performance assessments (2, 3, 4, 5, 7, 8, 10, and 11). To reduce the time required for the assessments, teachers and observers administered several of the activities during the same session. For example, the tangrams activity contained an assessment of both spatial (activity both spatial (activity 4) and logical-mathematical skill (activity 8). An art-based activity provided opportunities for assessment of spatial (activity 5) and linguistic intelligences (activity 10 and activity 11).

Project staff developed scoring rubrics for each performance activity after initial field tests of the assessments and teacher experiences during the previous year's administration. In addition, teachers and observers added to the rubric sheets as they assessed and evaluated student performance. For each activity, student performance and behavior were rated as "not evident or not observed" in a given setting (0), "evident" (a), or "extremely evident" (b). Teachers and external observers received training on administration of the assessments, with most having a year's experience with the assessments.

Table 2

Activities Used in the Alternative Assessments Categorized by Intelligence for Cohort 2

Activity Title	Type of Assessment
Spatial intelligence	
1. spatial checklist	teacher rating
2. Pablo	students construct 3-D animal from puzzle pieces
3. mechanical drain	students take drain apart, put it back together
4. tangrams	students manipulate puzzle pieces
5. artwork	students draw or paint a picture
Logical-mathematical intelligence	
6. math-logical checklist	teacher rating
7. bus activity	board game
8. math activity	students solve mathematical problems (including tangrams)
Linguistic intelligence	
9. linguistic checklist	teacher rating
10. storytelling activity	students tell a story using various objects (included storytelling based on art activity also)
11. pictorial prompt	students write, draw, or tell a story (included storytelling based on art activity also)
12. linguistic skills	observation during assessment battery
Interpersonal intelligence	
13. interpersonal checklist	teacher rating
14. interpersonal observations	ratings by teachers other than classroom teacher
15. interpersonal skills	observation during assessment battery

Data Collection and Analysis

Psychometric Analyses

Internal consistency and inter-rater agreement estimates were calculated using each student's data, while a random sample of students was administered specific performance measures after approximately four weeks to allow calculation of stability estimates. Traditional (Campbell & Fiske, 1967) and structural equation modeling (Jöreskog & Sörbom, 1989) approaches to multitrait-multimethod analysis were conducted to gather evidence of construct validity. The existence of gender and ethnic differences were investigated using multivariate analysis of variance (MANOVA) techniques. School was dropped as a dependent variable when initial investigations indicated there were no significant differences among schools.

Assessment of student linguistic-pictorial performance (activity 11) was accomplished through the use of up to two activities: oral skills and emergent writer. Students were allowed to write, describe, or both write and describe their pictorial work in order to allow for differences in student thinking style. Therefore, reliability calculations were performed separately for the emergent writer ratings and oral skills ratings.

Quantitative Student Outcomes

Research Questions

The first set of research questions address the traditional questions of reliability and validity of the tools used in identifying START students, the issues of potential ethnic and gender differences in performance on the instruments, and comparisons of the alternative assessments with traditional standardized assessments. Widespread use of alternative assessments without the study of their psychometric properties is indefensible.

The measurement of student change included quantitative and qualitative approaches. The nature of the students (highly at-risk because of economic or ethnic factors) raised the prospect that we would be able to affect achievement test score outcomes. Further, as explained in our description of treatment, the emphasis on instruction with a multicultural base and teaching to student strength was expected to positively affect attitude toward school subjects and student self-concept.

Specific research questions included:

4. What impact do the interventions have on student achievement?
    a. Do the interventions have different impacts on student achievement in four different academic subjects (vocabulary, reading, language arts, mathematics)?
    b. Do the interventions have different impacts on student achievement based on student ethnicity across the four subject areas (vocabulary, reading, language arts, mathematics)?
5. What impact do the interventions have upon student attitudes toward learning and education?
    a. Do the interventions have different impacts on student attitudes towards learning, teachers, language arts, and/or mathematics?

  - b. Do the interventions have different impacts on student attitudes based on student ethnicity toward learning, teachers, language arts, and/or mathematics?

6. Do Project START interventions have an effect on student self-concept?
   - a. Do the interventions have different impacts on specific areas of student self-concept?
   - b. Do the interventions have different impacts on student self-concept based upon student ethnicity?

Sample

Students were identified for cohort 1 as kindergarten and first-grade students during the spring of 1993. They were assessed on achievement, attitudes toward school, and self-concept in the fall of 1993; spring of 1994; and spring of 1995. At the conclusion of the study they were second and third-graders. The second cohort was identified in kindergarten in the spring of 1994. They were assessed in the fall of 1994 and spring of 1995. At the conclusion of the study they were completing first-grade (see Figure 2).

Instrumentation

Three instruments were used in this study to assess student outcomes. The Iowa Tests of Basic Skills were used to assess changes in basic skill achievement. The Arlin-Hills Attitude Surveys were used to assess changes in children's feelings about school, teachers, and learning. An instrument developed by the Instructional Objectives Exchange (IOX) (1972) was used to measure children's self-concept.

Iowa Tests of Basic Skills (ITBS)

The ITBS battery measures student growth in broadly defined skill areas. The basic purposes are to facilitate (a) within classroom decisions such as diagnosing strengths and weaknesses, and individualizing instruction; and (b) decisions external to the classroom such as identifying strengths and weaknesses of a group, and ascertaining the effectiveness of curricular or instructional modifications.

Internal consistency reliability estimates (K-R 20) reported by the authors ranged from .759 to .932 for Language (K-3), .769 to .910 for Mathematics (K-3), for .916 to .921 for Reading Comprehension (K-3), and .718 to .873 for Vocabulary (K-3).

	Spring 93	Fall 93	Spring 94	Fall 94	Spring 95
Cohort 1 Testing	Assessment	ITBS A-H SA	ITBS A-H SA		ITBS A-H SA
Treatment		⟶	⟶	⟶	⟶
Cohort 2 Testing			Assessment	ITBS A-H SA	ITBS A-H SA
Treatment				⟶	⟶

Note. A-H Arlin-Hills attitude surveys; SA: IOX self-appraisals.

Figure 2. Assessment, treatment, and testing schedule.

Arlin-Hills Attitude Surveys

Many educators note that high ability students may develop poor attitudes toward school if they are not taught in an appropriately stimulating environment (Clark, 1997; Tannenbaum, 1983). In order to address this issue, one of the questionnaires from the Arlin Hills Attitude Surveys was used to measure student attitudes toward learning processes (Arlin, 1976). This instrument assesses a student's perception of his or her degree of participation in classroom activities. The authors of this survey take the view that a student-centered classroom is a more positive learning environment than a teacher-centered environment. The 15-item instrument asks students to respond on a 4-point Likert response scale to items pertaining to attitudes about classroom activities such as the amount of homework they receive and the opportunities they have to work with friends throughout the day. Total scores range from 0 (low) to 60 (high) with a value of 30 or higher indicating a positive attitude. Standardization of the instrument took place in the spring of 1974 with over 13,000 students in grades 1 through 12 from a single southern state. Three levels of the instrument are available: primary for grades K-3, elementary for grades 4-6, and high school for grades 7-12. A description of the sample based on sex or racial/ethnic status was not provided in the manual (Arlin, 1976). The internal consistency reliability estimate reported by the authors for this survey was .90 across grades 1 through 12 ($n = 6{,}000$).

Self-Appraisal Inventory

The primary level of the Self Appraisal Inventory developed by the Instructional Objectives Exchange (1972) assesses self-concept. The test-retest (stability over two weeks) index was reported as .73, which is considered sufficiently high for affective group assessment. The content validity of the assessments was determined by expert reviews by educators, educational evaluators, and teachers.

Data Analysis

The research questions formulated for this portion of the study were answered using repeated measures ANOVAs. Separate ANOVAs were conducted for each type of outcome (e.g., ITBS, attitude surveys, and self-appraisals), and separate sets of analyses were performed for each cohort. Because of sample attrition across the project period and incomplete data, the sample for this portion of the study was substantially reduced from the previously reported sample making multivariate techniques inappropriate.

Gifted Program Participation

Research Question

Early identification and nurturing of talent is considered to be essential to the later recognition of gifted behaviors. In Project START, the expectation was that students who participated in the treatment condition would be identified in greater proportions than those not participating in the treatment conditions.

7. Are students who participate in Project START interventions referred and selected for entrance into the Charlotte-Mecklenberg Schools gifted program at a higher rate than control students?

An indicator of success for Project START was the percentage of treatment versus control group students referred and selected for entrance into the CMS academically gifted

(AG) program. Students were screened for participation in the AG program at the end of second-grade. Project START third-grade students were assessed in April of 1994, and second-grade students were screened in April of 1995. CMS staff in the office of gifted education provided referral and identification results for each treatment and control student.

Teacher Change

Research Questions

The ultimate change in students is dependent on teachers' willingness and ability to change their teaching behaviors, attitudes and/or values to a degree that allows the treatment to be strong and effective in classrooms. Many factors may influence that process. In our study of teacher change we attempted to assess the ways teachers responded to the philosophy of the program, the staff development activities and the coaching they received, and how this impacted their beliefs and behaviors.

8. How do teachers view the diversity of students in their classrooms after exposure to a model of individual differences such as that of MI theory?
9. How do teachers make meaning of the task of identifying student talent in diverse populations after exposure to MI theory?
10. How do teachers integrate notions of working with underserved students with a curriculum stressing language immersion, multiculturalism, manipulatives, and multiple talents and intelligences?
11. Are there specific developmental processes that teachers undergo as they attempt to implement a MI-based perspective on education, learning, and talent?

One pivotal component of Project START was the premise that persistent, focused, and prolonged staff development with a group of primary teachers could have a positive impact on their instructional practices in ways that positively affect the learning of high potential, low income, and/or minority students in those teachers' classrooms. It is therefore important to examine who the START teachers were, the nature of this facet of the START staff training intervention, means by which data were gathered and analyzed, the nature and degree of change resulting from the intervention, and factors that positively and negatively shaped the teacher change process. Understanding these elements not only facilitates interpretation of Project START outcomes, but also provides guidance for future research related to Project START in its design and goals.

Project START Teachers

Project START teachers taught in first-, second- or third-grade classrooms in schools that volunteered to participate in the project. Principals of the schools were requested by the Project START coordinator to select for participation in the project teachers who demonstrated strong skills in classroom instruction as well as an interest in the sort of staff training and change required to implement the project's goals. Principals also were asked to consider the importance of proportional participation of strong minority teachers in a project whose focus included improvement of schooling for minority children.

In some instances, principals asked for faculty volunteers for START participation. In others, teachers were drafted—occasionally against their will. In many instances, principals complied with the request to provide teachers with a level of skill suited to START goals. In a few instances, principals appointed START teachers with the hope that the project would enhance teachers' skills. Each participating school ultimately designated one first-grade, one second-grade and one third-grade teacher as START faculty.

Of the 48 START teachers (3 teachers in each participating schools), 66% were Caucasian, and 33% were African American. Teachers of first- and second-grade classes participated in staff interventions for all three years of the grant period. Third-grade teachers were phased in as START children moved into third-grade, and therefore took part in staff interventions only during the third (final) year of the funding period. All START teachers taught in "regular" classrooms in which there were approximately 25 students, six of whom were identified for START participation. START teachers were not released from any school or district mandates, duties, or expectations as a result of their status as START teachers, but rather their participation in START was "additive." START teachers averaged 14 years of teaching experience, with one having 31 years experience and seven being first-year teachers.

In the course of the study, four original START teachers were replaced due to illness or change of job assignment. In those instances, new START teachers were assisted by other START teachers in their school and the Project START coordinator in "coming aboard," but retrospective training provided these teachers was minimal in comparison with earlier training participated in by the teachers whom they replaced.

Data Collection and Analysis

Collecting Data Related to Teacher Change

Through the three-year duration of Project START, a variety of data gathering modes was employed by researchers to understand and study the response of participating teachers to START staff development interventions. Data were gathered through four primary means: (a) written surveys administered at the end of the first summer institute and near the end of the three years of interventions, in which teachers were asked to share their understanding of various START components and procedures, (b) persistent observation in START classrooms using structured observation protocols throughout the duration of the project, (c) use of focus group interviews with district resource teachers in programs for learners identified as gifted, and (d) formal and informal interaction with participating teachers and principals throughout the project.

Written surveys allowed testing of teachers' knowledge about START and its goals. Pre- and post-surveys administered to all START teachers were used for this purpose.

Use of classroom observations tested teachers' application of principles or knowledge how. All START teachers were formally observed by researchers at least three times per year throughout the study. Initially, observation data were gathered on the Classroom Practices Record (CPR) form VA2 (developed initially by the University of Connecticut of the NRC/GT for use in its Classroom Practices Study, later modified by the University of Virginia site of the NRC/GT for use in its Preservice Teacher Study, and finally further modified by the University of Virginia site for use in START). The modified CPR ultimately proved too general to elicit data needed to examine specific START issues and was replaced for years two and three of the study by the Project START Classroom Observation Checklist (see Appendix B for copies of both observation protocols).

Resource teachers for the district's gifted learners had general knowledge about START and its goals and had frequent interactions with all of the following groups: START teachers, all other teachers of grades one through three, students identified from START as eligible for district programs for the gifted, and non-START students identified as eligible for services for gifted learners. Their knowledge provided both an interesting "comparative" view of classrooms not available broadly to the researchers, as well as the only means available to determine how START students fared upon entry into district gifted programs in comparison with non-START peers. A two hour focus group interview with these teachers was conducted by researchers near the end of year three. Both written (individual) and oral (group interactive) responses were obtained.

During the three years of staff development sessions, general meetings, school visits, and chance interactions, researchers, principals, and teachers had many opportunities for exchange of ideas and feelings related to START. Researchers, acting as participant-observers, kept written records of these exchanges in the form of field notes.

Researchers included the Project START coordinator and Project START consultant who played a participant-observer role with the teachers throughout the three years of START. In addition, five graduate students from the University of Virginia also observed START classrooms and formally gathered data related to START instruction on at least ten two-day visits to the district.

Analyzing Data Related to Teacher Change

Data analysis regarding teacher change was on-going throughout the duration of Project START, guiding both subsequent staff training and data collection. The data analysis team was led by the project consultant and also included five graduate students on the staff of the NRC/GT. A constant comparison method of coding and theme emergence was employed with prose classroom observation records, teacher surveys, focus group notes, and researcher field notes. Member checks were used with participating teachers in two ways: (a) observations were followed by teacher/observer conferences in which observers reflected what they felt they were seeing during the observation and then discussed with the teacher her perceptions of the class, and (b) large group discussions between researchers and teacher participants provided an opportunity for researchers to share patterns that they felt they were seeing, and to ask the teachers to corroborate or modify the conclusions. Peer debriefing was employed in three ways: (a) periodic joint classroom observations conducted by the project consultant and project coordinator in which each took observational notes and then compared notes for similarities and differences following the shared observations, (b) formal sharing of patterns noted in separate observations by the project consultant and project coordinator twice yearly during the three-year project span, and (c) sessions among staff members at the NRC/GT comparing interpretations of START data.

Case Studies of Individual Students

In looking at the overall statistical results, we lose the understanding of what happens to particular children and the ways in which the dynamics of the treatment interact with the personality, environment, and ability of particular children. By selecting a small number of children who are differentially affected by treatment, we may gain insight into the ways in which the these factors interact in individual children.

Research Questions

12. In what ways do the interventions impact the lives of individual Project START students?

13. What are the effects of encouraging the parents of underserved students to become more involved in their children's education?
14. What are the differing dynamics in the lives of children judged to be successful and unsuccessful in Project START (according to changes in achievement test scores and teacher perceptions)?

Purpose of the Case Studies

High-risk, high-potential learners, such as those in Project START, often seem to fare poorly when assessed only via "traditional" standardized measures. Further, in-school measures do not allow a glimpse into connections between home and school that are important in any child's life, but that may have particular bearing on lives of children for whom poverty and/or cultural minority status are complicating factors.

Qualitative study, using observations and interviews, allows home-school connections to be explored, enables researchers to look for indicators of success or lack of success not likely to be reflected on standardized measures, and provides a mechanism for examining the impact of varied program components (singly and together) on the lives of students and families.

The case studies reported on in this section allowed for study in some depth and breadth of eight START learners. Further, they enabled researchers to determine factors that might have inhibited or promoted success of these learners in the project.

Sample

At the end of year two of Project START, all participating teachers were asked to nominate in writing a student in their class who appeared "successful" in START and one who appeared "unsuccessful." Further, they were asked to explain their reasons for the nominations. Teachers were told that four students from the "successful" category and four from the "unsuccessful" category would be selected by researchers for case studies to take place in the final year of the project. University of Virginia researchers selected eight students for case study based first on teacher comments and then on year one test score growth that seemed to corroborate the teachers' impression of success (that is, a student who showed marked growth on the ITBS during the first year of START) or non-success (that is, a student who made little, no, or negative growth on the ITBS during year one of START).

Four of the students selected were in third-grade during the case study span and four were in second. Selected students attended six different elementary schools, some rural and some suburban.

Researchers for the case study project were four graduate students at the University of Virginia with advanced training in gifted education. Instruction and guidance in qualitative methods were provided throughout the case study project for all researchers. Each researcher had major responsibility for developing case studies on two students. Researchers were blind to the successful/unsuccessful rating of students on whom case studies were developed.

Data Collection

Researchers each spent a minimum of six days on-site in the students' classrooms and conducting interviews with parents. Three site visits spread over three months allowed researchers to follow up with questions raised in early analysis of data as well as observation of students over time. Parents were notified of the case study project (they had given consent earlier in START for student participation in research related to the project) and their cooperation was sought. All parents contacted were supportive of the research and agreed to being interviewed for the case studies.

Initially, each researcher observed his/her case studies in his/her educational surroundings, including full-day observations, initial classroom teacher interviews, principal interviews, and interviews with the child's START teacher from the previous school year. Second round visits included interviews with target students, parent interviews, mentor interviews, family outreach coordinator interviews, classroom observation, and additional teacher interviews. Third round site visits included classroom observation and follow-up interviews based on data analysis to that point. At some point in the site visits, researchers also had an opportunity to observe mentorship orientation meetings, faculty meetings, special classes (e.g., English as a Second Language, resource classes for students identified as gifted), and school events (e.g., Christmas programs).

Initially, observations were made based loosely on project elements (e.g., multiple intelligences, language immersion, multiculturalism, and use of manipulatives in learning) as well as to determine a general profile of the learner as reflected in the class. Interviews began with a "grand tour" question (Spradley, 1979). Later, semi-structured interview protocols were developed and used, based on key project goals and questions that arose from first round interviews. Protocols differed somewhat depending on the role of the interviewee in START.

Researchers made extensive observation and field notes in the course of each site visit. All interviews were tape recorded, transcribed for analysis, and checked for accuracy and completeness by the researcher.

Following each field visit, researchers paired to debrief on field notes, observation notes, and transcripts. Researchers coded notes for recurrent patterns and ultimately themes. In addition, notes and transcripts were reviewed by a research coordinator to probe for additional questions, ambiguities, and themes. Each child's case was individually constructed by its primary researcher according to a case study analysis protocol developed to promote categorical consistency across cases. Categories included a description of school and home settings, a vignette of a typical school experience, the child's involvement in START (from observation as well as perspectives of parent, child, family outreach coordinator, mentor coordinator, principal, and other key players), curricular modifications observed and/or reported, evidence of the child's talent (strength) areas, and additional themes and impressions. Group debriefing sessions were held between each site visit to look for emergent common and disparate themes among the cases.

CHAPTER 4: Results and Discussion

The results and discussion chapter is divided along the same dimensions as the previous chapter on methodology and the findings will be presented in the same order and using the same framework. First we will discuss the findings related to the psychometric qualities of the MI instruments, then we will present the achievement data, the data relating to gifted identification, and the qualitative data on teachers and students.

Research Questions

1. Do the MI-based assessments show evidence of reliability and validity?
2. Are there ethnic and/or gender differences in student performance on the MI-based assessments?
3. How do the MI-based, alternative assessments compare to traditional, standardized assessments with respect to psychometric properties?

Psychometric Properties of Alternative Assessment Battery: Cohort 1

Reliability

For each of the subscales included in Table 1, Cronbach's alpha was calculated. Resulting values for alpha were acceptably large (Thorndike & Hagen, 1955): Logical-mathematical, .73; linguistic, .72; spatial, .74; interpersonal, .87.

Validity

Factor Analysis

Factor analysis was used to confirm the presence of four intelligences in the battery of activities. Principal factors extraction with varimax rotation was performed on the 13 activities using SPSS™ (Table 3). Variables were generally well-defined by the four resulting factors, with communality values ranging from .57 to .78. Overall, four factors accounted for 67.5% of the variance in student scores on the 13 activities.

Based on suggestions made by Comrey (1973), variables with loadings of .55 and above were interpreted. This rule of thumb was used because loadings of .55 and above are considered very good, accounting for 30% overlapping variance. The activities designed to assess linguistic and interpersonal intelligence (activities 9 through 13) all loaded onto the first factor; logical-mathematical activities (6, 7, and 8) loaded onto the second factor; and the five spatial activities loaded on the third (activities 1, 3, and 5) and fourth factors (2 and 4). The tangrams and Pablo activity were geometric in nature while the artwork, pump, and spatial activity were not necessarily geometric in design. The pump required manipulation of an object to see its functions and workings and its construction; the spatial checklist covered a wide variety of behaviors relating to spatial behaviors and the artwork was totally of the student's election in focus.

Construct Validity

Inter-item correlations appear in Table 4. High correlations among the teacher rating scales and the observation checklists (with the exception of the spatial scale) were noted.

Table 3

Factor Loadings, Communalities (η^2), and Percent of Variance for Principal Factors Extraction and Varimax Rotation for Performance-Based Assessment Activities—Cohort 1

Activity	F_1[a]	F_2	F_3	F_4	η^2
12-interpersonal checklist (I)[b]	.78	.38	.05	.18	.78
9-linguistic checklist (V)	.76	.33	.17	.13	.74
13-interpersonal observation (I)	.72	.38	.03	.22	.71
10-storytelling (V)	.71	-.10	.34	.04	.63
11-pictorial activity (V)	.56	.31	.39	-.01	.57
7-bus activity (M)	.15	.77	.10	.10	.64
8-math worksheet (M)	.25	.70	.24	.11	.63
6-math-logical checklist (M)	.40	.63	.35	.17	.71
3-pump activity (S)	.08	.13	.80	.16	.70
5-artwork activity (S)	.35	.31	.56	.24	.59
1-spatial checklist (S)	.32	.35	.55	.33	.63
4-tangrams activity (S)	.00	.22	.16	.82	.75
2-Pablo activity (S)	.26	.02	.18	.78	.70
Percent of variance	45.1	9.4	6.9	6.1	

[a] Factor labels:
F_1 linguistic - verbal; interpersonal
F_2 logical - mathematical
F_3 spatial - general
F_4 spatial - tangrams.

[b] Theoretical subscales in parentheses: (V) linguistic; (M) math; (S) spatial; (I) interpersonal.

Table 4

Interitem Correlations Among Assessment Activities and Checklists—Cohort 1

Act.	1(S)[a]	2(S)	3(S)	4(S)	5(S)	6(M)	7(M)	8(M)	9(V)	10(V)	11(V)	12(I)	13(I)
1	1.00[b]	.13	.11	.08	.17	.22	.14	.13	.19	.13	.13	.19	.17
2		1.00	.19	.30	.27	.28	.22	.23	.26	.23	.19	.23	.24
3			1.00	.41	.30	.27	.18	.20	.19	.13	.19	.20	.20
4				1.00	.45	.27	.20	.22	.19	.09	.20	.19	.23
5					1.00	.47	.33	.36	.42	.31	.54	.40	.44
6						1.00	.47	.53	.61	.33	.47	.58	.53
7							1.00	.38	.35	.22	.34	.38	.37
8								1.00	.41	.27	.42	.40	.40
9									1.00	.47	.54	.69	.63
10										1.00	.36	.37	.36
11											1.00	.50	.47
12												1.00	.78
13													1.00

[a] Subscales in parentheses: (V) linguistic; (M) math; (S) spatial; (I) interpersonal.
[b] For all coefficients $p < .01$.

A multitrait-multimethod matrix (Campbell & Fiske, 1967) appears in Table 5. Values in the validity diagonal are moderate for the math subscales and low for the language/linguistic subscales. Values in the heterotrait-heteromethod triangles are sufficiently low to provide evidence of discriminant validity, with the exception of the relatively high correlation between the ITBS language subscale and the math performance assessments. Unfortunately, this correlation exceeds the correlation between the ITBS language subscale and linguistic performance assessments. The ITBS language subscale also correlates more highly with the math teacher checklist than the linguistic teacher checklist.

Gender, Ethnic, SES, and School Differences

The entire sample of 1,813 students was used to investigate whether gender, SES, and school differences existed on the battery of assessment tasks and rating scales. Bonferroni's procedure was used to adjust for the influence of multiple analyses upon the study-wide alpha level ($\alpha = .003$). A two-way, between subjects ANOVA design with school (16 levels) and gender as independent variables revealed that a significant school effect was present for all four subscales, although the corresponding effect sizes were small (Table 6). Post hoc analysis utilizing the Tukey-b procedure at the .05 level indicate that school 4 had significantly lower scores on the math-logical and interpersonal subscales than over half of the other schools (although no pattern was apparent with respect to the other schools). For the linguistic subscale, schools 13, 12, and 16 scored significantly higher than a majority of the other schools. On the spatial subscale, schools 12 and 13 scored significantly higher than schools 14, 15, 4, 3, 2, and 7; school 15 scored significantly lower than 12 schools. No discernible, overarching pattern emerged from the analysis, suggesting inconsistency rather than systematic bias. With respect to the two significant gender effects, female students scored higher than male students on both the linguistic and interpersonal subscales, although the effect sizes were rather small.

Table 5

Multi-Trait Multi-Method Matrix—Cohort 1

		Performance Assessments			Teacher Checklists			ITBS	
		Math	Linguistic	Spatial	Math	Linguistic	Spatial	Math	Language
Performance Assessments	Math	1.00							
	Linguistic	.14*	1.00						
	Spatial	.11	.09	1.00					
Teacher Checklists	Math	.41**	.24**	.16*	1.00				
	Linguistic	.16*	.46**	.12	.36*	1.00			
	Spatial	.09	.20*	.47**	.50**	.27**	1.00		
ITBS	Math	.29**	.11*	.15*	.22**	.06	.09	1.00	
	Language	.25**	.16**	.06	.17**	.08	.07	.70**	1.00

*$p < .05$. **$p < .01$ (2-tailed).

Table 6

ANOVA Results for Assessment Subscales—Cohort 1

	School Effects			Gender Effects			School X Gender		
Subscale	F	p*	η^2	F	p^a	η^2	F	p^a	η^2
Math-logical	3.04	.001	.026	.01	.945	n/a	1.32	.182	n/a
Linguistic	6.16	.001	.051	12.22	.001	.007	1.43	.123	n/a
Spatial	7.28	.001	.060	4.12	.043	n/a	1.38	.150	n/a
Interpersonal	4.13	.001	.035	21.21	.001	.012	1.40	.140	n/a

*$p < .003$.

Because the number of African American and Caucasian students greatly exceeded the number of Hispanic American and Asian American students in the sample, students were randomly selected from the larger ethnic groups to equalize group size in the corresponding one-way ANOVA. Resulting ANOVAs with the Tukey-b post hoc procedure indicated that Asian students scored or were rated significantly higher than all other ethnic groups on all four subscales: math-logical, F(4, 149) = 12.57, $p < .0001$, $\eta^2 = .252$; linguistic, F(4, 142) = 4.49, $p < .002$, $\eta^2 = .112$; spatial, F(4, 79) = 13.42, $p < .0001$, $\eta^2 = .405$; and interpersonal, F(4, 157) = 5.01, $p < .001$, $\eta^2 = .113$. The only other difference between specific ethnic groups occurred on the spatial subscale—students classified ethnically as "other" by the school district (e.g., students of Asian Indian or mixed ethnicity) had significantly higher scores than Hispanic-American students.

Psychometric Properties of Alternative Assessment Battery: Cohort 2

Descriptive Statistics

The mean, standard deviation, kurtosis (peakness of the distribution), and skewness for each activity and scale appear in Table 7. Every activity and scale was significantly kurtotic, which is not surprising considering the use of a three-point Likert-type scale. Activities 3, 6, 7, and 10 had positively skewed distributions, as did the math-logical scale.

Reliability

Three types of reliability were estimated: inter-rater agreement, stability, and internal consistency. Inter-rater agreement was calculated for four of the performance assessments, with Cohen's kappa (Cohen, 1988) used as an unbiased measure of agreement. Agreement was sufficient for group assessment purposes: Emergent Writer activity, 90.3% agreement (kappa = .85); Oral Skills, 80.2% (kappa = .69); Storytelling 85.1% (kappa = .77); Visual-Spatial Skills 84.1% (kappa = .75).

Table 7

Descriptive Statistics for Assessment Activities and Scales—Cohort 2

Scale/Activity	Mean	SD	Kurtosis	Skewness
Spatial	.92	.49	-.69	.08
1. checklist	.81	.70	-.93	.27
2. Pablo	.95	.76	-1.24	.08
3. drain	.77	.78	-1.24	.44
4. tangrams	1.01	.77	-1.32	-.01
5. artwork	1.09	.75	-1.20	-.14
Math-Logical	.74	.61	-.91	.40
6. checklist	.77	.72	-1.03	.38
7. bus	.54	.75	-.52	.99
8. tangrams	.90	.86	-1.61	.19
Linguistic	.89	.56	-.82	.09
9. checklist	.92	.76	-1.27	.14
10. storytelling	.78	.76	-1.17	.38
11. writing prompt	1.00	.70	-.95	.00
12. skills	.85	.69	-.91	.21
Interpersonal	.86	.68	-1.04	.26
13. checklist	.88	.76	-1.23	.20
14. observation	.89	.75	-1.19	.18
15. skills	.82	.74	-1.14	.30

Stability calculations were computed for the following activities (Table 8): Pablo, mechanical drain, spatial tangrams, pictorial-visual, bus, math tangrams, and storytelling. Because of the restricted range of the assessment scores, percent agreement was calculated and also appears in Table 8. "Negative disagreement" is used to denote case in which the post-test score was higher than the pre-test score. Conversely, "positive disagreement" represents those cases where the post-test score was lower than the pre-test score. For every assessment activity, agreement was 50% or higher, and percent negative disagreement was higher than percent positive disagreement in every case with the exception of the math tangrams activity (in which they were equal). Considering the sample sizes, the similarity between percent agreement and the stability coefficients provides evidence that the concern about restricted range was unwarranted.

Cronbach's alpha was calculated as a measure of internal consistency. Alpha values ranged from a low of .66 for the spatial intelligence scale to .90 for the interpersonal scale, with values for the math-logical (.69) and linguistic scales (.77) also of considerable magnitude.

Validity

Evidence was gathered for two types of multitrait-multimethod analysis: a matrix using the performance assessments, teacher checklists, and Iowa Tests of Basic Skills; and confirmatory factor analysis using LISREL 7 (Jöreskog & Sörbom, 1989) with RAM notation (McArdle & McDonald, 1984). The matrix of correlations between the various components of the assessment battery and the ITBS appears in Table 9. The results are quite similar to that for the previous cohort, although many of the convergent validity correlations are relatively lower. As was also the case with the first cohort, the math checklists and ITBS subtests correlated with almost every other assessment and standardized test.

Table 8

Stability of the Performance Assessments

Task	*n*	Percent Negative Disagreement	Percent Agreement	Percent Positive Disagreement	Stability Coefficient[a]
Pablo	40	25.0	60.0	15.0	.48
drain	39	20.5	64.1	15.4	.55
spatial tangrams	46	26.1	52.2	21.8	.47
pictorial-visual	20	30.0	50.0	20.0	.47
bus	38	28.9	50.0	21.1	.60
math tangrams	45	22.2	55.6	22.2	.54
storytelling	37	35.1	43.2	21.6	.46

[a] All stability coefficients $p < .01$.

Table 9

Multi-Trait Multi-Method Matrix

		Performance Assessments			Teacher Checklists			ITBS	
		Math	Linguistic	Spatial	Math	Linguistic	Spatial	Math	Language
Performance Assessments	Math	1.00							
	Linguistic	.14	1.00						
	Spatial	.21**	.14	1.00					
Teacher Checklists	Math	.35**	.06	.13	1.00				
	Linguistic	.07	.39**	-.04	.41**	1.00			
	Spatial	.02	.10	.33**	.30**	.18**	1.00		
ITBS	Math	.37**	.16*	.16*	.24**	.17*	-.02	1.00	
	Language	.26	.04	.20**	.17*	.05	.10	.42**	1.00

$*p < .05$. $**p < .01$ (2-tailed).

Confirmatory Factor Analysis

Confirmatory factor analysis (CFA) allowed the estimation of fit for a variety of models based on Gardner's multiple intelligences, hypothesized method effects, and combinations of both MI theory and method effects. Table 10 contains the results of the CFA for the null model (no hierarchical structure), unitary model (one latent variable), theoretical model (four uncorrelated latent variables), a hierarchical model (four correlated latent variables), method effect models (latent variables representing type of assessment and not theoretical relationships), and hierarchical-method effect models. Based upon the numerous chi-square and goodness-of-fit measures, the hierarchical model appears to have the best fit to the data. Correlations among the four latent, first-order variables in the hierarchical model were quite high, ranging from .566 to .795 with a mean of .677. These correlations provide considerable evidence that the structure is characterized by only low to moderate distinction among the first-order factors.

Table 10

Results and Goodness-of-Fit for Confirmatory Factor Analyses

Model	df	X^2	*p*	GFI	AGFI	BBI	Δdf	ΔX^2	*p*
Null	107	7085.27	< .001	.345	.168				
Unitary	93	1970.84	< .001	.783	.683	.722	14	5114.43	< .001
Theoretical	98	3591.23	< .001	.769	.679	.493	9	3494.04	< .001
Hierarchical	91	1209.05	< .001	.880	.821	.829	16	5876.22	< .001
Method effect 1	94	3002.03	< .001	.769	.666	.576	13	4083.24	< .001
Method effect 2	96	2715.90	< .001	.790	.703	.617	11	4369.37	< .001
Method effect 3	96	2663.83	< .001	.812	.734	.624	11	4421.44	< .001
Method effect 1 with theoretical model	88	4301.67	< .001	.706	.545	.393	19	2783.60	< .001
Method effect 2 with theoretical model	90	4341.05	< .001	.674	.507	.387	17	2744.22	< .001
Method effect 3 with theoretical model	87	3932.82	< .001	.759	.624	.445	20	3152.45	< .001
Method effect 1 with hierarchical model	79	1792.28	< .001	.863	.764	.747	28	5292.99	< .001
Method effect 2 with hierarchical model	81	1971.41	< .001	.849	.746	.722	26	5113.86	< .001
Method effect 3 with hierarchical model	81	1604.92	< .001	.874	.789	.773	26	5480.35	< .001

Note. GFI=Goodness-of-Fit Index
AGFI=Adjusted Goodness-of-Fit Index
BBI=Bentler-Bonett Index.

Gender and Racial Differences

A two-way multivariate analysis of variance (MANOVA) was performed using the four MI subscale scores as dependent variables and student gender and race as independent variables. Due to the relatively small number of Asian American and Hispanic students ($n = 56$), they were not included in the analysis. Another 192 students were excluded due to missing data, resulting in a final sample size of 829 students for the MANOVA. The gender by race interaction was not significant (Wilks'Λ = .993, F(4, 822) = 1.37, $p = .244$, power = .43, canonical correlation = .081), while the race (Wilks'Λ = .886, F(4, 822) = 26.47, $p < .001$, power = 1.00, canonical correlation = .338) and gender main effects (Wilks'Λ = .968, F(4, 822) = 6.85, $p < .001$, power = .99, canonical correlation = .180) were each further analyzed with a two-group descriptive discriminant analysis (Huberty, 1994). Results of the discriminant analyses are presented in Table 11.

The canonical correlations and F-to-remove statistics indicate that the significant gender effect was small and attributable to higher female ratings on the interpersonal and linguistic scales. The ethnicity effect was of more moderate magnitude and attributable to the higher ratings of Caucasian students on each scale, especially the math-logical scale.

Quantitative Student Outcomes

In reporting the student outcome data, we have presented separate analyses for the achievement test data, the attitudes toward school and learning data, and the self-concept data. Further, we were interested in the potential differential affects on the ethnic minority groups so we have included that dimension in each analysis.

Table 11

Means and Discriminant Analysis F-to-Remove for Gender and Ethnicity Effects Upon MI-based Assessment Scales

	Female ($n = 473$)		Male ($n = 461$)			Caucasian ($n = 234$)		African American ($n = 595$)		
SDQII Scale	mean	sd	mean	sd	F-to-Remove[a]	mean	sd	mean	sd	F-to-Remove[b]
Math-logical	.76	.61	.73	.61	1.79	1.06	.58	.62	.58	34.59
Linguistic	.96	.55	.83	.56	4.25	1.11	.50	.82	.56	3.63
Spatial	.93	.48	.93	.50	2.90	1.12	.46	.85	.48	1.64
Interpersonal	.94	.69	.76	.64	9.94	1.05	.63	.79	.68	1.96

Note. SDQII=Self Description Questionnaire-II.

[a] Canonical r = .167.

[b] Canonical r = .337.

Research Questions

4. What impact do the interventions have on student achievement?
    a. Do the interventions have different impacts on student achievement in four different academic subjects (vocabulary, reading, language arts, mathematics)?
    b. Do the interventions have different impacts on student achievement based on student ethnicity across the four subject areas (vocabulary, reading, language arts, mathematics)?
5. What impact do the interventions have upon student attitudes toward learning and education?
    a. Do the interventions have different impacts on student attitudes towards learning, teachers, language arts, and/or mathematics?
    b. Do the interventions have different impacts on student attitudes based on student ethnicity toward learning, teachers, language arts, and/or mathematics?
6. Do Project START interventions have an effect on student self-concept?
    a. Do the interventions have different impacts on specific areas of student self-concept?
    b. Do the interventions have different impacts on student self-concept based upon student ethnicity?

Impact on Student Achievement

Although four subtests of the ITBS were used to determine the intervention effects on achievement, because of an inadequate number of students taking each subtest each testing period, multivariate analyses were not employed. Instead, four separate 3 x 2 x 4 univariate repeated measures analysis of variance (ANOVAs) were used to analyze student performance on each subtest at all three grade levels separately (first, second, third). Students in the first-grade were administered the ITBS subtests twice. Students in the second and third-grades were each administered the ITBS subtests three times over a two year period.

For research question 4a, the independent variable was treatment group (Project START intervention without mentor component, intervention with mentor component, and control group).

Table 12 presents the means and standard deviations for the ITBS subtests by treatment group for grade one. As can be seen from the table, all three groups' means increased across the two testing periods. For the mentor group, the largest increase occurred on the mathematics subtest with an increase of .74 grade equivalent (GE) units. The vocabulary subtest increased by .64 GE units, the reading comprehension subtest by .66 GE units, and the language arts subtest by .64 GE units. For the intervention group without the mentor, all ITBS subtest scores also increased across the two time periods with the largest increase occurring in the reading comprehension subtest, .87 GE units. The mathematics subtest mean score increased .77 GE units, with the language arts subtest mean score increasing slightly less than a half of GE unit (.48). The vocabulary subtest had the least mean increase, .29 GE units. For the control group's performance, all ITBS subtest scores increased over time with the mathematics subtest experiencing the largest gain, .87 GE units, followed closely by the reading comprehension subtest, .84 GE units. A .58 GE units mean increase occurred on the language arts subtest followed by a .32 GE unit increase on the vocabulary subtest. There were no consistent pattern of mean increases across the three treatment groups; no one group consistently had the largest mean increases nor did one group consistently have the smallest mean increases across any of the subtest areas.

The results of the four univariate ANOVAs (Table 13) for grade 1 indicate that there was a statistically significant Retest main effect for the vocabulary, reading comprehension, language arts, and mathematics subtests (to control for Type I error, Bonferroni's procedure was used employing a significance level of .0125). Regardless of the treatment group students were assigned to, there was a significant difference in the ITBS subtests' scores across the two testing periods, implying that no one particular intervention had an effect on student achievement.

Means and standard deviations for the ITBS subtest scores for grade 2 are given in Table 14. Again, regardless of student group assignment, all three groups experienced increases in mean scores across all four subtests. For the intervention group assigned a mentor, mean score increases ranged from 1.23 GE units on the mathematics subtest to 1.65 GE units on the language arts subtest. For the intervention group without a mentor, mean score increases ranged from 1.07 GE units on the mathematics subtest to 1.68 GE units on the language arts subtest. Although the control group's mean scores increased on all four subtests, the group's increases were less than the other two groups with the largest mean increase occurring on the language arts subtest (1.54) and the least mean increase occurring on the vocabulary subtest (.70). In general, the largest mean increases occurred in the two treatment groups with the control group experiencing the least mean increase for all four subtests.

Table 12

Means (Sds) for ITBS Subtests* by Treatment Group for Grade 1

	ITBS Subtest							
	Vocabulary		Reading		Language Arts		Mathematics	
	T1	T2	T1	T2	T1	T2	T1	T2**
Mentor	0.99	1.63	1.18	1.84	1.22	1.86	.90	1.64
n = 55	(.76)	(.83)	(.54)	(.63)	(.61)	(.85)	(.50)	(.65)
No Mentor	1.38	1.67	1.29	2.16	1.23	1.71	1.12	1.89
n = 54	(.93)	(.81)	(.46)	(.69)	(.56)	(.74)	(.47)	(.70)
Control	1.55	1.87	1.28	2.12	1.34	1.92	1.24	2.11
n = 52	(.99)	(.98)	(.60)	(.69)	(.70)	(.91)	(.60)	(1.46)

*GE Scores.
**Time 1 and Time 2.

Table 13

ANOVA Results for Treatment Effects Upon Achievement for Grade 1

Source of Variation	df	MS	*F*	*p*	Power
VOCABULARY					
Between Subject					
Treatment	2	4.38	3.48	.033	.644
Error	158	1.26			
Within Subject					
Retest	1	14.01	46.00	.000	1.000
Retest x Treatment	2	.98	3.23	.042	.610
Error	158	.30			
READING					
Between Subject					
Treatment	2	1.23	2.10	.127	.425
Error	135	.59			
Within Subject					
Retest	1	42.97	283.39	.000	1.000
Retest x Treatment	2	.26	1.72	.183	.356
Error	135	.15			
LANGUAGE ARTS					
Between Subject					
Treatment	2	.80	.97	.382	.216
Error	158	.82			
Within Subject					
Retest	1	29.04	110.07	.000	1.000
Retest x Treatment	2	.44	1.65	.195	.344
Error	158	.26			
MATHEMATICS					
Between Subject					
Treatment	2	4.33	4.99	.008	.807
Error	15	.87			
Within Subject					
Retest	1	50.57	128.08	.000	1.000
Retest x Treatment	2	.12	.30	.743	.097
Error	157	.39			

Table 14

Mean (Sds) for ITBS Subtests* by Treatment Group for Grade 2

	ITBS Subtest											
	Vocabulary			Reading			Language			Mathematics		
	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3**
Mentor	.97	1.78	2.50	1.27	1.97	2.68	1.26	1.84	2.91	1.25	2.02	2.48
(*n* = 42)	(.66)	(.84)	(.92)	(.26)	(.41)	(.71)	(.62)	(.76)	(.97)	(.46)	(.82)	(.87)
No Mentor	1.18	1.83	2.65	1.49	2.02	2.91	1.31	2.00	2.99	1.26	2.23	2.33
(*n* = 41)	(.81)	(.92)	(.96)	(.48)	(.66)	(.84)	(.47)	(.76)	(.89)	(.61)	(.71)	(.72)
Control	1.42	1.61	2.12	1.43	2.13	2.64	1.33	2.16	2.87	1.23	1.94	2.44
(*n* = 22)	(.96)	(.70)	(1.05)	(.41)	(.69)	(.64)	(.52)	(1.04)	(1.01)	(.63)	(.71)	(.95)

*GE Scores.
**Time 1, Time 2, and Time 3.

The results of the four univariate ANOVAs for grade 2 are presented in Table 15. A significant Retest main effect for student achievement was indicated for the reading comprehension subtest, language arts subtest, and the mathematics subtest. Follow-up Tukey-b post hoc procedures were used to determine which means across the three testing administrations were significantly different. (For the vocabulary subtest, a significant retest x treatment interaction effect was indicated.) Results of the Tukey-b procedure (Table 16) indicated that all three means were statistically different for all three subtests: reading comprehension, language arts, and mathematics. As can be seen from Table 15, for each test administration, regardless of treatment group, there was an increase in the mean achievement. Figure 3 displays each treatment group's profile across the three testing periods. The patterns indicate that the intervention groups' profiles had the same shape (i.e., parallel). This finding implies that all simple interaction effects were zero. However, the control group's profile compared to the other two groups departs significantly from parallelism. That is, for the first administration the control group had the highest mean score on the vocabulary subtest, however by the third test administration the control group had the lowest mean score. These trends suggest that the control groups' means were significantly different from both the intervention groups' means.

Table 15

ANOVA Results for Treatment Effects on Achievement by Treatment Group for Grade 2

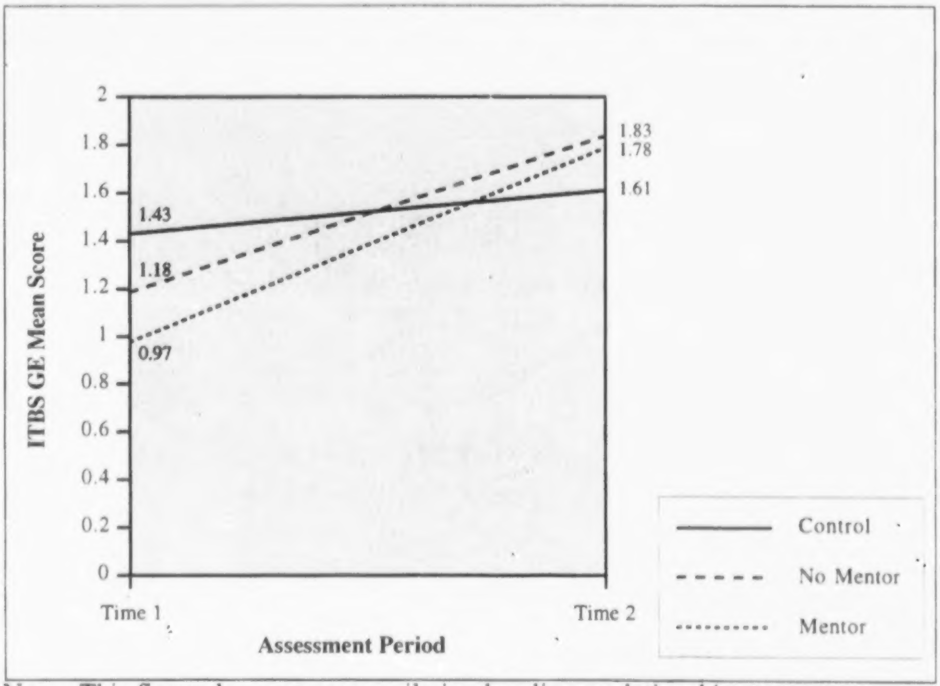
Source of Variation	df	MS	*F*	*p*	Power
VOCABULARY					
Between Subject					
Treatment	2	.85	.59	.558	.145
Error	102	1.46			
Within Subject					
Retest	2	36.59	91.20	.000	1.000
Retest x Treatment	4	1.55	3.85	.005	.892
Error	102	.40			
READING					
Between Subject					
Treatment	2	.47	.64	.532	.152
Error	60	.74			
Within Subject					
Retest	2	25.98	132.44	.000	1.000
Retest x Treatment	4	.21	1.06	.381	.325
Error	120	.20			
LANGUAGE ARTS					
Between Subject					
Treatment	2	.43	.37	.693	.108
Error	98	1.16			
Within Subject					
Retest	2	62.02	175.56	.000	1.000
Retest x Treatment	4	.25	.71	.586	.227
Error	196	.35			
MATHEMATICS					
Between Subject					
Treatment	2	.07	.07	.930	.060
Error	58	1.00			
Within Subject					
Retest	2	19.70	71.37	.000	1.000
Retest x Treatment	4	.29	1.04	.388	.321
Error	116	.28			

Table 16

Tukey-b Post Hoc Comparisons for Grade 2 Achievement

	Null Hypothesis	q-statistic
READING	M3-M1=0	24.39*
	M3-M2=0	13.36*
	M2-M1=0	11.04*
LANGUAGE ARTS	M3-M1=0	27.75*
	M3-M2=0	11.47*
	M2-M1=0	16.27*
MATHEMATICS	M3-M1=0	16.81*
	M3-M2=0	12.51*
	M2-M1=0	4.29*

*$p < .05$.



Note: This figure does not necessarily imply a linear relationship.

Figure 3. Vocabulary Retest x Treatment interaction effect—Grade 2 achievement.

Table 17 displays the means and standard deviations for grade 3. Again, regardless of treatment group, for each of the ITBS subtests there were mean increases. Overall, grade 3 experienced the largest mean increases for each subtest when compared to the other two grade levels with the largest increases occurring in the language arts and mathematics subtests.

Results of the four univariate repeated ANOVAs for grade 3 can be found in Table 18. Again, the Retest main effect was statistically significant for each of the four ITBS subtests. Follow-up post hoc Tukey-b procedures are displayed in Table 19. As can be seen from the table, all mean scores were significantly different from each other.

Table 17

Means (Sds) of ITBS Subtests* by Treatment Group for Grade 3

	ITBS Subtest											
	Vocabulary			Reading			Language			Mathematics		
	T1**	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3
Mentor	2.20	2.76	3.69	2.40	2.97	3.68	2.09	3.43	3.88	1.89	2.51	4.20
($n = 41$)	(.86)	(.79)	(.99)	(.73)	(.92)	(.92)	(.88)	(1.20)	(1.18)	(.85)	(.70)	(1.28)
No Mentor	2.01	2.63	3.61	2.16	2.75	3.74	2.15	3.40	4.01	2.03	2.42	3.97
($n = 40$)	(.97)	(.95)	(1.31)	(.61)	(.84)	(1.26)	(.85)	(1.33)	(1.94)	(.71)	(.95)	(1.50)
Control	2.00	2.74	3.59	2.10	2.84	3.32	1.71	3.09	3.50	1.83	2.37	4.05
($n = 26$)	(.86)	(.92)	(1.54)	(.57)	(.81)	(1.25)	(.68)	(1.17)	(1.71)	(.91)	(.95)	(1.40)

*GE Scores.
**Time1 and Time 2.

Table 18

ANOVA Results for Treatment Effects Upon Achievement for Grade 3

Source of Variation	df	MS	*F*	*p*	Power
VOCABULARY					
Between Subject					
Treatment	2	.72	.31	.734	.098
Error	103	2.33			
Within Subject					
Retest	2	63.08	148.65	.000	1.000
Retest x Treatment	4	.03	.08	.998	.066
Error	206	.42			
READING					
Between Subject					
Treatment	2	2.19	1.26	.288	.268
Error	97	1.74			
Within Subject					
Retest	2	41.63	116.59	.000	1.000
Retest x Treatment	4	.42	1.17	.324	.364
Error	194	.36			
LANGUAGE ARTS					
Between Subject					
Treatment	2	4.93	1.49	.230	.311
Error	98	3.31			
Within Subject					
Retest	2	86.38	113.91	.000	1.000
Retest x Treatment	4	.16	.21	.933	.094
Error	198	.76			
MATHEMATICS					
Between Subject					
Treatment	2	2.41	.06	.939	.059
Error	99	.15			
Within Subject					
Retest	2	119.35	251.41	.000	1.000
Retest x Treatment	4	.34	.71	.588	.227
Error	198	.47			

Table 19

Tukey-b Post Hoc Comparisons for Grade 3 Achievement

	Null Hypothesis	q-statistic*
VOCABULARY	M3-M2=0	15.13**
	M3-M1=0	24.73**
	M2-M1=0	9.60**
READING	M3-M2=0	11.33**
	M3-M1=0	22.22**
	M2-M1=0	10.88**
LANGUAGE ARTS	M3-M2=0	5.45**
	M3-M1=0	20.95**
	M2-M1=0	15.50**
MATHEMATICS	M3-M2=0	23.81**
	M3-M1=0	31.24**
	M2-M1=0	7.43**

$^*q_{.05} = 3.64$. $^{**}p < .05$.

Impact on Student Achievement by Treatment Group and Student Race

Attrition rates associated with transient, low SES students impacted the number of students remaining in Project START at the end of the project. Consequently, statistical analyses to address research question 4b could not be run due to insufficient numbers of students in each cell.

Impact on Student Attitudes by Treatment Group

Student responses to the Arlin-Hills Attitude Surveys were analyzed using 3 x 2 x 4 univariate repeated measures ANOVAs. The dependent variable for each of the ANOVAs was student scores on each of Arlin-Hills Surveys: Learning, Teachers, Language Arts, Mathematics. For one set of analyses the independent variable was the treatment group (mentor, no mentor, control). The second set of analyses could not be completed due to attrition rates of the low SES students. Statistical analyses were completed separately for the first cohort (n = 109) and the second cohort (n = 109). Again univariate analyses as opposed to multivariate analyses were used due to inadequate sample sizes within each treatment group.

Treatment Group Effects

Means and standard deviations for cohort 1 are presented in Table 20. For the learning, teacher, and language arts surveys all three treatment groups experienced a decrease in mean scores and an increase in score variability. For the mathematics survey, the treatment group assigned the mentor produced a decrease in the mean score across the testing period. However, the other two groups, intervention with no mentor and control, experienced an increase in mean scores with the control group showing the largest mean increase, 2.52 points.

Results from the ANOVAs for cohort 1, student attitudes are displayed in Table 21 (research question 5a). Again, controlling for type I error, only the learning and language arts surveys resulted in significant mean differences. Regardless of group membership, on the attitudes toward learning survey, the decrease in mean scores from the first test administration to the second resulted in a statistically significant difference. On the attitudes towards language arts survey, although the two intervention groups' mean scores decreased and the control groups' mean score increased across the two testing periods, the difference between the overall mean scores from the testing periods resulted in statistically significant differences.

Table 20

Means (Sds) for Arlin-Hills Attitude Surveys by Treatment Group for Cohort 1

	Attitude Subareas							
	Learning		Teachers		Language Arts		Mathematics	
	T1	T2	T1	T2	T1	T2	T1	T2*
Mentor	19.38	17.18	29.98	29.30	27.14	30.80	28.72	26.14
n = 50	(6.28)	(7.11)	(9.06)	(9.94)	(11.16)	(10.18)	(9.73)	(11.63)
No Mentor	21.14	19.67	33.08	30.22	27.75	32.86	29.86	32.08
n = 36	(7.47)	(7.95)	(9.01)	(11.19)	(9.27)	(10.26)	(11.62)	(10.84)
Control	23.57	21.30	32.00	30.22	28.17	32.39	28.87	31.39
n = 23	(4.70)	(9.48)	(8.11)	(8.67)	(8.75)	(9.38)	(8.53)	(9.23)

*Time 1 and Time 2.

Table 21

ANOVAs of Retest Effect for Arlin-Hills Attitude Surveys Upon Student Attitude for Cohort 1

Source of Variation	df	MS	*F*	*p*	Power
LEARNING					
Between Subject					
Treatment	2	177.79	3.13	.047	.594
Error	158	56.87			
Within Subject					
Retest	1	362.52	5.84	.017	.671
Retest x Treatment	2	2.62	.04	.958	.056
Error	158	62.09			
LANGUAGE ARTS					
Between Subject					
Treatment	2	82.25	.74	.480	.174
Error	168	111.45			
Within Subject					
Retest	1	688.00	7.58	.007	.781
Retest x Treatment	2	8.45	.09	.911	.064
Error	168	90.79			

Cohort 2 means and standard deviations are displayed in Table 22. It can be seen that for the teacher and language arts surveys, mean scores increased for all three treatment groups, with the intervention group without a mentor having the largest mean increases on both surveys. However, the opposite pattern occurred in the learning survey; all three treatment groups' mean scores decreased again with the intervention with no mentor group experiencing the largest decrease. On the mathematics survey, both intervention groups' mean scores increased with the control group's mean score slightly decreasing.

Table 22

Means (Sds) for Arlin-Hills Attitude Surveys by Treatment Group for Cohort 2

	Attitude Subareas							
	Learning		Teachers		Language Arts		Mathematics	
	T1	T2	T1	T2	T1	T2	T1	T2*
Mentor	20.28	18.00	20.06	26.06	22.11	29.93	26.13	26.61
n = 35	(6.97)	(8.24)	(4.53)	(12.01)	(5.38)	(9.98)	(8.69)	(12.11)
No Mentor	22.26	19.77	22.37	30.32	22.15	30.50	27.74	30.71
n = 38	(7.44)	(8.36)	(5.71)	(8.87)	(4.67)	(11.91)	(9.20)	(9.97)
Control	21.69	19.37	22.47	26.94	21.36	26.46	25.68	24.97
n = 36	(7.34)	(7.84)	(4.21)	(12.04)	(4.62)	(13.49)	(8.49)	(12.32)

*Time 1 and Time 2.

Results of the univariate ANOVAs for cohort 2 student attitudes are presented in Table 23. Significant differences were found in the learning, teachers, and language arts surveys. There were no significant differences on the mathematics survey. The Retest main effect was significant across all the surveys, indicating again that regardless of group, the differences in scores across the two testing administrations were statistically different.

Treatment and Race Effects Upon Attitudes

As indicated earlier, attrition rates of low SES students across the project period prevented statistical analysis for these breakdowns (research question 5b).

Impact on Student Self-Concept

Means and standard deviations for cohorts 1 and 2 for the self-concept areas are presented in Tables 24 and 25.

Table 23

ANOVAs of Treatment Effect for Arlin-Hills Attitude Surveys Upon Student Attitude for Cohort 2

Source of Variation	*df*	MS	*F*	*p*	Power
LEARNING					
Between Subject					
Treatment	2	62.53	.82	.444	.187
Error	99	76.40			
Within Subject					
Retest	1	283.64	6.64	.011	.723
Retest x Treatment	2	.21	.00	.995	.051
Error	99	42.70			
TEACHERS					
Between Subject					
Treatment	2	196.69	2.46	.090	.485
Error	106	79.89			
Within Subject					
Retest	1	2052.15	31.35	.000	1.000
Retest x Treatment	2	56.21	.86	.427	.194
Error	106	65.47			
LANGUAGE ARTS					
Between Subject					
Tetest	2	105.41	1.17	.314	.252
Error	103	90.02			
Within Subject					
Retest	1	2576.82	37.26	.000	1.000
Retest x Treatment	2	46.02	.67	.516	.159
Error	103	69.19			

Table 24

Means (Sds) of Student Self-Concepts by Treatment Group for Cohort 1

	Self-Concept Areas							
	General		Family		Peer		School	
	T1*	T2	T1	T2	T1	T2	T1	T2
Mentor	5.44	5.22	4.63	4.64	5.04	4.69	5.58	5.52
n = 73	(1.34)	(1.26)	(1.83)	(1.38)	(1.82)	(1.40)	(1.63)	(1.38)
No Mentor	5.08	5.04	4.41	4.62	4.31	4.78	5.45	5.49
n = 74	(1.43)	(1.25)	(1.28)	(1.48)	(1.31)	(1.44)	(1.29)	(1.24)
Control	5.10	5.07	4.61	4.37	4.39	4.20	5.51	5.20
n = 41	(1.20)	(1.19)	(1.38)	(1.24)	(1.16)	(1.31)	(1.36)	(1.29)

Table 25

Means (Sds) of Student Self-Concept by Treatment Group for Cohort 2

	Self-Concept Areas							
	General		Family		Peer		School	
	T1*	T2	T1	T2	T1	T2	T1	T2
Mentor	5.06	5.60	4.44	4.96	4.67	4.98	5.32	5.51
n = 52	(1.45)	(1.36)	(1.41)	(1.40)	(1.59)	(1.50)	(1.46)	(1.25)
No Mentor	5.43	5.20	4.78	4.47	4.02	4.74	4.98	5.39
n = 54	(1.41)	(1.17)	(1.45)	(1.46)	(1.50)	(1.33)	(1.59)	(1.32)
Control	5.00	5.27	4.57	4.53	4.49	4.40	5.35	5.14
n = 49	(1.32)	(1.02)	(1.47)	(1.57)	(1.59)	(1.38)	(1.58)	(1.23)

*Time 1 and Time 2.

Cohort 1

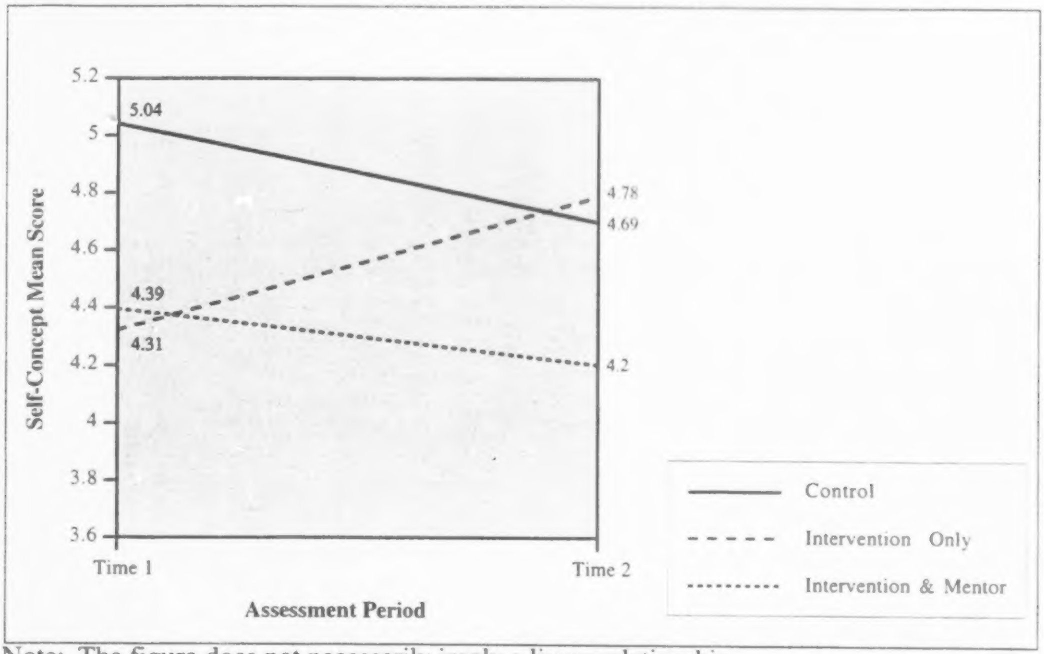
Based on the means in Table 24 for the general self-concept area, all three groups experienced decreases in mean scores across assessment periods. The mentor treatment group experienced the largest mean decrease, .22 score units. The intervention group without a mentor and the control group experienced only slight decreases in mean scores across the assessment periods. The intervention group without a mentor experienced the largest increase across the time period, .21 score units. The control group's mean scores across the two time periods decreased by .24 score units. For the peer self-concept area, only the intervention group with no mentor had a mean score increase, .47 units. The intervention group with a mentor experienced the largest mean decrease going from 5.04 to

Table 26

ANOVA Results for Treatment Upon Student Self-Concept for Cohort 1

Source of Variation	*df*	MS	*F*	*p*	Power
PEER					
Between Subject					
Treatment (T)	2	6.98	2.68	.071	.526
Error	203	2.61			
Within Subject					
Retest (R)	1	.13	.08	.773	.671
R x T	2	8.00	5.00	.008*	.056
Error	203	1.60			

*$p < .01$.



Note: The figure does not necessarily imply a linear relationship.

Figure 4. Peer self-concept Retest x Treatment interaction effect—Cohort 1.

BEST COPY AVAILABLE

4.69 across the two assessment periods. For the control group, the mean decrease was .19 units, going from 4.39 to 4.20 from time 1 to time 2. In the school self-concept area, the intervention group without a mentor experienced a slight mean score increase (.04 score units), where the two other groups experienced a mean score decrease. The control group's mean score decreased the most going from 5.51 to 5.20. The intervention group with a mentor treatment group's mean score decreased by .06 score units across the two assessment periods, 5.58 to 5.52. In the area of family self-concept, the intervention group assigned a mentor essentially remained the same (4.63 versus 4.64). The intervention group without a mentor gained a minimal amount across the time period, 4.41 versus 4.62. The control group experienced a slight decrease in the area of family self-concept across the time period, 4.61 versus 4.37.

Cohort 2

Patterns similar to cohort 1 were observed for cohort 2. However, there were some slight variations (Table 25). In the general self-concept area, only the intervention group with no mentor experienced a mean score decrease, 5.43 to 5.20. The intervention group with a mentor experienced the largest mean score increase, 5.06 to 5.60, across the two assessment times. For the control group, the mean score increase was .27 score units, going from 5.00 to 5.27. Similar patterns in the general self-concept area were found in the family self-concept area, with the only mean score increases occurring with the intervention group with a mentor and mean score decreases occurring with both the intervention group without a mentor and the control group. For the peer self-concept area, both intervention groups experienced mean score increases while the control group's mean score decreased. The largest mean increase occurred with the no mentor intervention group, .72 score units. The control group's mean scores across the two testing periods decreased by .09 score units, 4.49 to 4.40. In the school self-concept area, both intervention groups again experienced increases in mean scores and the control group experienced a mean score decrease. The intervention group with no mentor experienced the largest mean increase, 4.98 to 5.39, with the mentor intervention group's mean score increasing from 5.32 to 5.51. The control group's mean score decreased across the two assessment times going from 5.35 to 5.14.

For both cohort 1 and 2, it should be remembered that only slight mean score increases or decreases are probably due to testing error rather than actual differences in scores. Therefore, all results are essentially nil.

Treatment Effects

To address research question 6a on the impact of self-concept for both cohort 1 and cohort 2, separately, univariate repeated measures analysis of variance procedures were conducted. Table 26 presents the results for cohort 1. As can be seen from the table, only the peer self-concept area had any significant differences. Specifically, there was a significant retest x treatment interaction effect. Figure 4 displays the profiles of the three groups. From the figure, one can see that the treatment group with no mentor differed significantly from both the control group and the intervention group with the mentor; a parallel pattern was found between the intervention group with a mentor and the control group across both testing periods. However, the profile for the intervention without a mentor group did not follow the same trends and consequently as can be seen from the figure intersected both the other two groups.

The were no statistically significant differences found in any of the self-concept areas for cohort 2.

Gifted Program Participation

The analysis of the questions relating to the influences of Project START on the ultimate identification and placement of START students rested on analysis of the referral and placement data comparing students in the treatment conditions to students in the control group.

Research Question

7. Are students who participate in Project START interventions referred and selected for entrance into the Charlotte-Mecklenberg Schools gifted program at a higher rate than control students?

Referral and placement statistics for third- (April, 1994) and second-grade (April, 1995) students appear in Tables 27 and 28, respectively. In 1994, referral and selection rates were higher for treatment students (55% and 33%, respectively) than for control group students (39% and 14%, respectively), with a similar pattern in 1995: 61% referral, 29% placement for treatment and 30% referral, 9% placement for control students.

Table 27

Referral and Selection Information for Project START Students Currently in Grade 3

	Treatment Group			Control Group		
School	*n*	# referred	# selected	*n*	#referred	# selected
1	9	3	2	6	0	0
2	6	6	2	4	2	0
3	8	6	6	7	5	1
4	8	2	2	3	1	1
5	7	4	3	2	0	0
6	9	4	3	1	0	0
7	8	1	1	3	1	1
8	8	6	3	7	3	1
9	9	7	5	2	2	2
10	7	2	0	2	0	0
11	5	3	3	1	0	0
12	8	6	4	3	3	1
13	6	4	1	3	1	0
14	8	3	3	5	1	0
15	6	4	0	4	2	1
16	6	4	1	4	1	0
Total	118	65	39	57	22	8

Table 28

Referral and Selection Information for Project START Students Currently in Grade 2

	Treatment Group			Control Group		
School	*n*	# referred	# selected	*n*	#referred	# selected
1	7	3	0	2	2	0
2	6	5	5	2	2	0
3	6	3	1	8	4	1
4	7	4	2	4	1	0
5	7	4	2	0	0	0
6	9	7	2	7	1	0
7	5	0	0	1	0	0
8	5	4	0	0	0	0
9	7	5	4	4	4	4
10	7	1	0	7	1	0
11	5	4	3	1	0	0
12	6	1	1	1	1	0
13	8	6	3	3	0	0
14	6	2	2	6	0	0
15	6	6	2	2	2	0
16	7	4	3	6	0	0
Total	104	63	30	54	16	5

Nature and Degree of Change Observed in START Teachers

With the disappointing achievement test data, we looked to the information on the nature and degree of change in START teacher beliefs, values, and attitudes for insight into the ways in which we had or had not made fundamental changes in the delivery of curriculum to gifted students. This qualitative analysis was used to help understand the impact of the staff development process, the experiences the teachers had in implementing the curriculum, and the ways in which they did or did not integrate the precepts of the START interventions in their classrooms.

Research Questions

8. How do teachers view the diversity of students in their classrooms after exposure to a model of individual differences such as that of MI theory?
9. How do teachers make meaning of the task of identifying student talent in diverse populations after exposure to MI theory?
10. How do teachers integrate notions of working with underserved students with a curriculum stressing language immersion, multiculturalism, manipulatives, and multiple talents and intelligences?
11. Are there specific developmental processes that teachers undergo as they attempt to implement a MI-based perspective on education, learning, and talent?

In the Beginning

As START began, the Charlotte-Mecklenburg Schools had a strongly articulated, superintendent-led mandate to raise student standardized test scores. Not only was this goal highly publicized in the local press, but teacher salaries were directly tied to a school's reaching its prescribed (by district formula) benchmark goals. Teachers talked often about principals who "disappeared overnight" because test scores in their schools did not improve. Whether reality bore out this teacher perception was not tested in the study, but the perception was evident, widespread, and persistent.

Impacts of what became a (if not "the") prime district goal on teacher change will be discussed in more detail later in this section. At the outset, however, this test-improvement focus appeared as at least one evident contributor to primary grade classrooms that were often teacher-centered and drill-and-skill-driven. In a typical classroom, all students practiced a single prescribed list of skills for much of prime instructional time. Teachers checked off skills which had been covered. Variance in student readiness, interest, and learning style were seldom issues of concern, because, as one teacher said, "Nobody cares if we take a student with low skills and make progress with him. Even really good progress that most teachers would get excited about doesn't count. If we can't get him up to the standards by test time, we might as well not have tried. Nobody cares that he grew. It just doesn't count."

Project START premises included beliefs that low economic and/or minority learners: (a) often fare poorly in traditional classrooms that center on part-to-whole, drill-and-practice learning, (b) would develop deeper understandings of key concepts and principles in classrooms that provide learners ample, concrete opportunity to construct meaning or make sense of ideas through high level thinking, manipulation of materials, and student-centered discussion of ideas, (c) would feel a greater sense of belonging, motivation and success in classrooms that acknowledge and build on their particular learning strengths, and (d) would relate better to learning attached to their home and community environments via materials, simulations, stories, illustrations, and patterns that seem familiar. District primary classrooms did not exemplify these beliefs as the study began. In early observations, for example, teachers were noted to set up a spatial learning station, a kinesthetic activity, and a verbal task—but the three were often neither focused on any particular class learning goal nor on mutually reinforcing ideas. At worst, a spatial activity was putting together a jigsaw puzzle in a first-grade classroom with no reference to classroom content. At best, a kinesthetic activity might be lining up to rehearse cardinal numbers rather than writing them on paper.

START staff developers were then left with a choice—affirm the teacher-centered, skill-and-drill approach to learning that many teachers felt compelled to pursue, and encourage application of START elements in that setting, or work with the teachers to develop a different approach to instruction in general, then incorporate both START elements and required skills standards in the new approach. The latter seemed more likely to address the needs of START learners and to lead toward improved classroom practice in general. The former seemed destined to give license to something other than best teaching practice and to sentence non-traditional learners to the sort of fragmented and pointless exercises that often defeat children such as those identified for START. While the latter approach was believed to be clearly more defensible based on current pedagogical theory, it was bulkier from a research standpoint as it caused staff developers and participating teachers to have to "go backwards and forwards" in their thinking simultaneously.

As reflected in the START Classroom Observation Checklist, two additional goals were added to the START staff development agenda in an attempt to help teachers develop student-centered learning environments. Along with understanding and applying START elements of multiple intelligences, manipulatives, language immersion, and multiculturalism, teachers were encouraged to develop their ability to base instruction on key concepts and principles of meaningful content (including balancing skills and ideas, balancing content and process, integrating high level thinking into tasks, and stressing meaningful use of skills and understandings by students), and to modify instruction based on on-going assessment of student interest, readiness, and learning profile.

Assessment of START teacher growth, then, will be examined first in those three arenas that teachers studied and were encouraged to apply throughout Project START.

Teacher Growth in Key Components of START

In analyzing teacher-growth data, it appears clear that the three-year project span was adequate for teachers to develop a considerable bank of insight about constructing a Project START classroom (e.g., concept-based, differentiated in response to learner need, using manipulatives, language immersion, multicultural approaches, and multiple intelligences as a way of understanding students and designing learning). The duration of the project was less adequate in allowing time for complete translation of understanding into classroom practice.

Written surveys indicated that even early in the project, teachers could "talk the talk" of START. "I will need to relate my learning experiences to the real world of the child so each child will see why we learned what we learned." "I will need to use flexible groups in my class so that my students can work in their strength areas which will be different at different times." "I need to remember that all students are unique and have special strengths, that all students want to be successful and belong, and that they can be successful and belong if I give them learning opportunities best suited to their learning strengths and styles."

Formal and informal conversations with teachers indicated consistent knowledge about developing a Project START classroom. More difficult was translating the knowledge about into knowledge how—making the leap into classroom practice. For example, one teacher noted, "I understand it now—in my head. What I suppose I still have to do is get it out of my head and into my lesson plans." In a post-observation conference, a teacher who had taught a single, fact-based, worksheet-oriented lesson to all students was asked if she could think of anything she might do if she were to reteach the lesson in order to address the START elements. She replied, "I could have thought about what big idea the students need to learn about matter (in science). Then I could have developed an activity which would let different students explore that idea using different intelligence areas."

Nonetheless, by the end of year 3 most first- and second-grade teachers were making the transfer of principles into classroom practice and reported and were observed using manipulatives, language immersion, multicultural approaches, and multiple intelligence avenues during START observations.

In capsule, first- and second-grade START teachers were exhibiting more small group work, more active learning, more use of learning centers, use of manipulatives in mathematics, extensive use of multicultural literature, more discussions with students, and more writing and more kinds of writing among students. Facets of START elements less prevalent in participating classrooms by the end of year 3 were: use of manipulatives to help students understand big ideas in science and language arts, multiculturalism reflected in exploration of varied perspectives on events, student oral storytelling, and consistent use of high-level thinking in tasks and discussions, and group or individual projects centered on "real" problems.

Growth in Multiple Intelligence Instruction

Because of START's particular focus on identifying and teaching to a child's intelligence, teacher understanding and translation of MI theory merits more extensive comment here. Probably the most powerful impact of MI on teachers was attitudinal rather than instructional. MI provided (a) a language for talking about student strengths beyond decoding, encoding, and computation, and (b) made concrete the hope that students brought with them "invisible" but important abilities. Early on, there was a clear sense in teacher talk (formal and informal) that students who could not read and write according to traditional conceptions of intelligence could not be "smart." "I don't see how these (first-grade) students could be identified as gifted (e.g., for START). They can't read." However, START teachers were willing relatively quickly to test the belief that students could be smart in ways other than traditional language and math, and were willing to be a part of looking for such evidences. Noted one teacher, "I just feel more positive about them (START learners) now. I just have this belief that there's something locked up inside and I can help unlock it." MI became a metaphor for untapped possibilities. If someone said the START children were smart, then maybe they were. Especially with relatively compliant START children (see discussion of "successful" and "unsuccessful" START learners), MI helped teachers see them as "special in a positive way." The affirmative teacher attitude along with implicit special status in the classroom likely resulted in powerful environmental differences for the START learners.

In regard to *teaching* via multiple intelligences, most first- and second-grade teachers seemed able, by project's end, to develop and use multiple activities for a single task goal. That led to more active, student-centered classrooms as well as to increased interaction between teacher and individual students or small groups of students.

Two specific patterns emerged that indicated barriers to more extensive and richer use of multiple intelligences as a guide in planning and teaching. First, teachers largely persisted in a belief that to be fair, they must have all students participate in all activities. While this stance makes classroom management simpler, and may allow all students exposure to development of a range of intelligences within themselves, it falls short in two areas. One, if all students must take part in all activities, it is difficult to fashion student-specific routes to learning which enable particular individuals to use intelligence strengths to buoy areas of weakness (e.g., to use spatial strengths as a consistent avenue to improve word recognition or comprehension). Two, if all students must be participants in all activities, it is difficult to commend a differentiated approach to multiple intelligences whereby students with special talents in an intelligence area are guided in continued escalation of the intelligence area. One-size-fits-all multiple intelligence instruction runs the same liability as most one-size-fits-all instruction—that is, underestimating the capacity of talent and projecting expectations at a sort of middling level of quality.

Requiring highly talented students to work at the same level of functioning as less talented students in a given talent area was another liability. In many of the materials commercially available to teachers on multiple intelligences in the classroom, MI is treated much like another learning style theory. Allowing a child to work kinesthetically is said to mean, for example, having the child move while he learns—for instance, act out a story, dance an idea, etc. If kinesthetic talent were an intelligence rather than a learning style, and if the teacher's role were seen as developing that intelligence in a talented child, a much more complex set of questions emerges. The teacher would need to understand what constitutes kinesthetic talent, at least a descriptive scope and sequence of kinesthetic skills, descriptors of quality for talent development in that area, and how to assess readiness and progress in kinesthetics itself. To date, it would appear that few educators and theoreticians have attempted to address these issues as they relate to MI. Lacking insight into the answers, teachers shift the focus from the talent area (e.g., kinesthetics) to the school-based task (e.g., understanding the story just completed)—and the talented child is likely not to have his kinesthetic talent developed, but rather acts out stories and dancing ideas instead of learning how to be a better actor or dancer. While researchers, administrators, and teachers with Project START confronted these issues together, implications of the issues are huge in scope and point to the need for research and reflection beyond the scope of this project.

Growth in Concept-Based Instruction

For many START teachers, virtually all instruction at the outset of the study was skill-based (e.g., emphasis on mastery of cardinal and ordinal numbers, computation in isolation, defining terms). In the relatively rare instances where a more content-directed approach to teaching and learning was employed, the content tended to center on holidays (e.g., Thanksgiving) or potentially isolated topics (e.g., penguins). Concern for high-risk learners who appear in special need of context-rich and power-producing learning experiences, START staff developers began late in year one of the project to work with teachers on developing concepts that might help students (and teachers) organize ideas, make connections among them, and derive principles that govern the way things work in the world. Classroom observations indicated that without a meaning-based approach to learning, multiple instruction would remain game-like and tangential to the sort of sense-making necessary for high potential, high risk learners.

With guidance through START staff development, and working in grade-level teams, teachers first developed banks of concepts which seemed to them potentially useful in binding together and illuminating previously isolated skills and spheres of study at a given grade level. Next, they created "location maps"—a sort of web designed to help them test a given concept for use in a range of classroom explorations. Once a grade-level team selected a key concept and an initial investigation to which it might be applied, teachers in that team worked to develop guiding principles or generalizations that they felt would be important for students to understand and apply related to the selected area of investigation. They then worked to develop a sequence of content and process that would acquaint students with the concept and generalizations and come to understand through exploration how the generalizations worked in the area being explored. Teachers then examined their draft lessons for effective ways to ensure that multiple intelligences, language immersion, use of manipulatives, and multiculturalism were reflected in student learning options, and to ensure that key skills specified by the district as essential were incorporated in meaningful ways in the lessons. Ultimately, the teachers developed "exploration tubs" in which they placed the concept-based, multiple intelligence-focused lessons, related study materials such as books and computer disks, and instructions for (or samples of) activities in which students might engage to learn about the concepts and make sense of the generalizations in meaningful applications. The tubs were plastic boxes with lids that could be circulated among START teachers throughout the district. Not only did it allow economy of time through shared planning and materials production, but teachers were encouraged to add to the tubs as they used them, thus promoting in-process revision and extension of instructional plans.

Not surprisingly, this was a major departure in teacher planning from worksheet-oriented, single-skill lessons. However, it was too much of a leap for some teachers to make fully during the three years of Project START. Teachers generally found it difficult to "locate" the meaning which undergirded the things they taught from the district curriculum and thus frustrating to determine key concepts implicit in the materials. It was then difficult to come to understand a concept as a professional might define it. For example, the concept of "culture" was initially seen as talking about neighborhoods and holidays—both related to the concept in a way, but certainly falling short of a more dimensional sense of the elements of culture and their interrelationships. Developing generalizations and sequencing them in a way which seemed optimally meaningful for student exploration was also difficult (which, for example, led to an inclination to introduce definitions in the middle or end of a study, if at all). It was also difficult for many of the teachers to think about how to use the district-mandated skills in context rather than teaching them in isolation.

The range of teacher growth in the area of concept-based instruction was evident. Among subscribers was a teacher who remarked with visible enthusiasm, "This is great! It's the first time I've really understood what I've been teaching in the six years I've taught! Can you imagine how much more likely my students are to understand it now?" Several Project START teachers created bulletin boards or charts designed to help students reflect on the elements of a concept, and systematically referred to the bulletin boards during class discussions and activities, linking what was studied during reading time, for example, with something examined during science, and using the concept and its elements as a link. Several teachers talked excitedly over time about beginning with one concept and becoming aware that it was evolving into another. One teacher, for example, began by working with her students on "systems" in mathematics and science. In time, she found that her understanding and that of her students was clarified by moving between the concepts of "systems" and "interdependence."

Less robust growth in use of concept-based instruction were teachers who developed concept-based charts or bulletin board, but found it difficult to help students link what they were doing to the more concrete manifestations of the concept. Some teachers simply did not use the concept-based materials. From this group, a large number cited as a reason the pressure they felt to "cover the skills" mandated by the district and a fear that more meaning-based instruction would result in lower test scores. Others said they planned to use the newly-developed concept-based materials in the future, but didn't "have the time to make the changes it would take to do that right now."

Growth in Differentiation of Instruction

All data indicate that Project START teachers as a group were consistently more likely to differentiate instruction for students based on interests and/or learning profile via use of a multiple intelligence orientation by the end of the project period than at its beginning. In this regard, multiple intelligence materials became a lever for differentiation in that they caused teachers to think about learner differences rather than about similarities only. For example, there was evidence in most of the later observations (especially among first- and second-grade teachers) that students had choices of ways to express learning, and that students who needed a more active orientation to making sense of ideas were more "acceptable" to teachers than was the case in the early months of Project START. Defined in this way, differentiated instruction was far more evident in Project START classes than in non-START classes of the same grade by the conclusion of the three years.

Defined as modification of instruction in relationship to student readiness (as opposed to student interest and/or learning profile), differentiation was probably the least successful element of the Project START staff development interventions—at least as it related to advanced learners. For struggling learners, there was a sort of differentiation by default more than by design because students' lack of progress made necessary a recycling of explanation and practice. Differentiation for advanced learners was seldom noted in observational data. There seemed to be two key reasons for the absence. First, teachers were already overwhelmed with calls for instructional modifications in the forms of multiple intelligence, concept-based instruction, use of manipulatives, multiculturalism, and language immersion. The suggestion that, in addition, they might want to think about instruction that would be challenging to a child who was advanced in, for example, kinesthetics or mathematics was simply too great an expectation. Said one teacher, "I know he's having to wait on a lot of the others in math, but there's just no time in my life, no room in my head, to do anything about it now." A second common reluctance to differentiate instruction for readiness level was a belief among primary grade teachers that they would be perceived as unfair if all students did not do all work. "They (students) get angry if they don't go to all the centers. They're always asking, 'How come I don't get to do that?'" "They get really upset if they don't do what everybody else does. It's like (they think) I don't like them or want to punish them or something." Teachers did not feel that they could assist students in conceiving the classroom in a light other than a place where everybody does everything. Rather the teachers perceived that a strongly held student value prescribed the course of action they must take.

While the majority of Project START teachers evidenced growth in translating one or more elements of the instructional model into their classrooms, there is little evidence to suggest that most did so in a coherent and integrated way. That is, few teachers appeared to make connections in their thinking between a concept and a student's need to make sense of ideas, multiple intelligences, and multicultural approaches. It appears that the elements remained discrete skills to be demonstrated by teacher participants rather than closely coupled responses to students with a particular learning profile. Both the lock-step approach to teaching all students alike and the lack of coherence perceived among the Project START curricular elements were likely aggravated by a major district-level focus on standardized testing. This phenomenon will be discussed later.

Teachers as Identifiers of Talent

Use of multiple intelligence approach to teaching had two evident impacts on Project START teachers as talent scouts. While most teachers readily accepted the promise that identification could reveal the student's undiscovered capacity, they were less likely to accept that a category of talent suggested by the identification was a child's best or only talent area. Teachers insisted that they discovered strength areas not delineated by identification results. This happened, they asserted, as they provided each child with a range of multiple intelligence activities rather than restricting options to strength areas suggested in a student's identification profile. Consistently teachers told stories of a child who was identified in one area but who "is really better" in another. For example, one first-grade teacher described a boy labeled as verbally talented. "He really likes the spatial activities better, though. Always picks those. Really shows his ability there."

The study did not provide clear evidence of why teachers consistently reported "alternative strengths." Teachers' own preferences and strengths may have impacted task quality for better or worse. Novelty of task may have been a draw for students. Whatever the reason, staff development sessions and teacher interviews were frequently marked with teacher stories about students' newly discovered talent areas. The tone and excitement of the stories was positive.

Second, a number of Project START teachers became firm champions of and advocates for one or more of their START students. They seemed to realize an opportunity to foster the students' development by supporting placement in the district's program for gifted learners. The language of multiple intelligences often figured into teacher descriptions of these students. "She has logical and verbal strengths. She's good with personal (skills), too. She'll be good (in the gifted program). One teacher reflected when three of her START students were offered placement in a magnet school for gifted learners. "I'm so proud. I mean, it'll be a loss to the (research) project and all if they go, but just think what it's going to mean for them."

Once subscribers to the notion of great possibilities in START students, teachers often became sources of additional "evidence" of student strength. For students on the receiving end of these "discoveries," classroom environments doubtless became more positive.

Case Studies of "Successful" and "Unsuccessful" START Students

The overall achievement, self-concept, and attitude changes were not significantly different for START students; however, teachers and project coordinators were quick to point to individual students who were profoundly affected by the project intervention. They were also quick to question the identification process in particular cases and the ineffectiveness of the interventions with certain students. It is apparent that the lives of students were affected individually by the interventions we selected. The case study methodology yielded insights into the ways in which Project START had differential effects depending on the circumstances of students, families, and the teachers who worked with the students.

Research Questions

12. In what ways do the interventions impact the lives of individual Project START students?
13. What are the effects of encouraging the parents of underserved students to become more involved in their children's education?
14. What are the differing dynamics in the lives of children judged to be successful and unsuccessful in Project START (according to changes in achievement test scores and teacher perceptions)?

At the end of the second year of Project START, all participating teachers were asked to nominate in writing one student from their class who appeared "successful" and one who appeared "unsuccessful" and to provide a rationale for their choice. Four students were selected for the successful category based on teacher nomination and relatively larger gains on the standard achievement measures. Four unsuccessful students were selected based on teacher nomination and relatively less success on standardized assessments. Staff from the NRC/GT spent a minimum of six days on-site. They spent time in the students' classrooms and they interviewed current and previous Project START teachers, students, students' parents (or primary caretaker), mentors, and the family outreach coordinators. Earlier sections of this document described the analysis in detail.

Following each field visit, researchers paired to debrief on field notes, observation notes, and transcripts. Researchers coded notes for recurrent patterns and ultimately themes. In addition, notes and transcripts were reviewed by a research coordinator to probe for additional questions, ambiguities, and themes. Each child's case was individually constructed by its primary researcher according to a case study analysis protocol developed to promote categorical consistency across cases. Categories included a description of school and home settings, a vignette of a typical school experience, the child's involvement in Project START (from observation as well as perspectives of parent, child, family outreach coordinator, mentor coordinator, principal, and other key players), curricular modifications observed and/or reported, evidence of the child's talent (strength) areas, and additional themes and impressions. Group debriefing sessions were held between each site visit to look for emergent common and disparate themes among the cases.

The eight cases themselves are lengthy, and are thus not included in total in this report. Included here are synopses of two "successful" and two "unsuccessful" students to provide readers with some insight into both the students involved and the case studies themselves. The synopses are followed by a discussion of key themes from all eight cases.

Two "Successful" Project START Students

Charelle

Charelle is an African American student in the third-grade, on free lunch, homeless for much of the year, and much loved and supported by both parents. Her mother is a housekeeper in a local hospital. Her father "flips burgers" (her words) at a fast-food restaurant. Now in housing in a different school zone, Charelle still attends the school in which she began, because her mother makes the long bus ride with Charelle, continuing on to her own job via public transportation. Charelle is often as much as an hour late for class because of the extended bus ride, but when she arrives in her classroom, she becomes immediately absorbed in her school work. Charelle's teacher feels the long ride seems worthwhile to Charelle's parents because the school has been nurturing to the family, and that Project START "may have been the icing on the cake (that kept them coming)."

Charelle seems to be hungry, not so much for food as for knowledge. She often asks for extra schoolwork to do at home. Her current teacher calls her ". . . a joy. I feel lucky to have her in my class. There are few children intrinsically motivated like Charelle . . . . She's a real big ham. She would act out anything. She's just kind of bright and bubbly and effervescent and gregarious . . . . She writes. She loves to tell stories. She's a good leader in a group . . . not as a forceful leader, but she coaches, like, 'well, maybe we should do this.' She's blown the doors off math in here. I have her well into fourth-grade math." Charelle's second-grade teacher echoes, "She's very talented in writing and reading. She is very creative, good in art, good in all subjects . . . ." The teacher points out a piece of Charelle's art work that is permanently displayed in the school corridor.

The researcher who observed and interviewed in Charelle's case study calls her "a ray of sunshine undimmed by her circumstances." Her mother, who seems quite aware of ideas underlying this multiple intelligence program, explains that "(Charelle) has very good writing skills. I'd say literature is a pretty smart area for her." Her father says, "She does what she sets out to do, and then when she's finished, she'll show us what she's done."

In her school, there is no Project START mentorship. Her parents say they have had difficulty coming to Family Outreach activities, but talked with enthusiasm about plans to attend an upcoming Project START activity that would include bowling. The Family Outreach Coordinator, however, says, "I think Project START has had an impact on Charelle and her family because they have been to most of our meetings."

Charelle's second-grade teacher appeared to be a subscriber to the instructional methods central to Project START. Her classroom is flexible, center-based, and gives evidence of a multiple intelligence emphasis. "I knew that Charelle was a good writer, so I really used her a lot in a lot of linguistic activities . . . . I was making sure the children were in the seven intelligences, and this was in the lessons we did, and because of that, I think Charelle was able to choose activities that were very good (for her)."

Her third-grade teacher, a first-year teacher who became both a teacher and a Project START teacher just prior to the opening of school, seemed less clear about the nature of START instruction, "I keep my START students with my assistant during math because they are faster-paced. They do problem-solving with multiplication, single and double digits, while my other students are figuring out how much a quarter is." Nonetheless, this teacher, too, establishes a flexible environment. "I think (Charelle) loves, enjoys, feels comfortable being in this environment as opposed to an environment that says, 'sit straight up, speak only when you are spoken to, school is not fun' type of environment." Students in this third-grade room were allowed to investigate and speak freely. "That's what I want her to do, is be able to take a risk in an environment were she knows I won't crush her. So I feel like this is a little bit of a START environment, where (students) are able to do things they wouldn't be able to do in the normal classroom environment, where the learning is self-directed."

The school principal seemed less than effusive about Project START, noting that it "seems to be of benefit to the children." She added that being asked by program administrators to keep a log of contacts with parents was "a little much."

Throughout the duration of the case studies, Charelle's attitude and work remained sterling. Absenteeism became more of a factor with each month. Teachers expressed doubt about whether she would finish the year.

William

William is a third-grade African American student whose observer called him "verbally calm, cool and collected, and behaviorally explosive." He does not receive free lunch at school. His mother works hard—extra hours around Christmas "to buy presents for me." William is very attuned to his mother's work schedule and whether or not she will be home when he is. His father appears to be largely an absentee in his life. His current teacher reports that his two key male role models are uncles—one who plays basketball in Europe, and one recently incarcerated. He has no siblings. On the other hand, the school's Family Outreach Coordinator reports that, "His parents always come to our Family Outreach meetings. I mean, when they can . . . . You can't always expect that they drop whatever (they are doing) and come have lunch with the kids all the time . . . . They are very supportive of him and Project START."

Sometimes when William's mother comes to school, however, she is defensive and accusatory with his teachers. On one occasion, she became verbally abusive with his third-grade teacher. The teacher asked her to bring her son to school the next morning so they could all talk about the issue at hand. She didn't show up. She wrote a note and said she couldn't make it. She apologized and made a total turnaround. A similar incident happened with William's second-grade teacher, who asked William's mother to come the following day to talk with the principal. During that meeting, his mother was "the picture of support and understanding." During the case study interviews, she was supportive, cooperative, and complimentary of his teachers and the school.

William plays pranks on adults and peers alike in ways that are humorous and generally well received. He and a friend used fake blood to stage a series of "cuts" that resulted in a series of students being sent to the nurse, until she became wise to the origin of the injuries. He introduced the interviewer to his cousin, telling her all about the things they did together as relatives. Later, of course, the interviewer discovered the two boys were not related. In her field notes, she commented, "William can be very convincing."

His third-grade teacher suggests, "He's very good at relating to children and adults. His verbal expression skills are incredible. I'm sure you've talked with him, and you know that. He's not intimidated in any way by adults, and with questions, he's very free to give his response, his feelings, and his opinions." "He's sensitive to the feelings of the other students. He's careful not to be pompous about being identified for START. Some students let the others know when they get to do something special. He doesn't brag or anything like that." At another point, however, the teacher notes (and observer confirms) that he has trouble with peer relationships. "I've had to move several students (away from William) . . . . He's getting into trouble more often . . . . I've even had a phone call from a mother requesting that (her daughter) be moved away from him."

William does not like to be touched. "When students bump into him in the busy classroom, he takes it as a challenge and defends himself." At the end of the case study observations, when the observer told him she would not be coming back to his class again, William became combative and would not make eye contact or verbal contact with her for the rest of the day.

William has had a mentor for each of his two years in Project START. Both mentors report enjoying him. Last year's mentor tells of being sure to gauge his mood before she began a session with him, "I would hang back and let him take the initiative when he was ready." She says this helped their working relationship and that she tried communicating this strategy to his teacher, "to help smooth whatever was happening in his life at the time." His third-grade mentor says he was "forewarned" about William by last year's mentor. He gives his two mentees "a little time to goof around, but they don't take advantage of it." William likes working on the stage in the cafeteria, so the mentor often plans activities that can be done there.

Both William's second- and third-grade teachers lead classrooms that seem immersed in kindness. Both use center-based classrooms, with his third-grade teacher appearing to emphasize a multiple intelligence approach more than his second-grade teacher, who seems to stress affect over content and process. William's second-grade teacher uses puppets to help the students express themselves, ensures that dioramas on the wall represent many cultures, and talked about channeling his potential "bossiness to good use" by naming him director of a play the class was going to perform. "It was a time he really shined in her class."

His third-grade teacher has many learning centers around the room that reflect a variety of "intelligences," although she says she does not consciously plan for every intelligence on a daily basis. She has even moved out her teacher's desk so there will be more space for student exploration. She talks in terms of William's strength areas. "He's very good about communicating his needs. He doesn't stumble over his words or anything . . . . He drew the three-dimensional figures (in the geometry mural) himself. I didn't help him with that . . . . He was extremely fluid and adept with this (spatial) activity. He knows how to make three-dimensional rectangles, prisms, cubes and things like that . . . . A lot of the students have no idea how to make a cube using toothpicks." This teacher also worked to use a concept-based approach to teaching her third-graders—for example, geometry lessons based on congruence and non-congruence.

William sees himself in terms of academic strengths too, although he seems unaware of a vocabulary directly related to multiple intelligences. "I really like math. (We were working with) congruent figures . . . . I drew a figure (a castle) and then had to make my space figures congruent . . . and I had a pyramid, rectangular prism, and a cube in my castle." William does not refer to his behavior problems. His observer called him "a complicated little individual with an exceptional mind . . . a man of many faces."

Two "Unsuccessful" Project START Students

Tev

Tev's family immigrated to the United States from Cambodia 10 years ago, with the aid of a sponsor in this country—a relationship that has been maintained for their decade of residence here. Tev is a second-grader with two working parents and a sister born the year after Tev's parents arrived in Charlotte. He is not on free lunch, and the school's mentor coordinator reports that while she did not know where the parents worked, she felt the family was "financially sound." The school psychologist says both children appear well adjusted and that the school has received no requests from the family for any sort of assistance. Tev has an extended family in the immediate area—uncles who are also immigrants from Cambodia. His grandmother comes to this country for extended visits from time to time.

At home, the telephone is answered in English. All other family communication is conducted in Cambodian. Tev's father is very attentive to his son and Tev reads to his father every evening. There is an active Family Outreach program at Tev's school, and his parents have attended every activity that it has sponsored. However, his father struggles with English, seems to have little idea about the meaning of Project START, and expresses a concern to know from the interviewer whether his son is doing well in school. He knows his son is good with computers and hopes someday to be able to buy one for him to use at home. School personnel see Tev's father as a "silent partner." No one in the school appears to have attempted to use a translator to facilitate communication with the family.

Tev's school is 78% African American and 10% ESL (English as a Second Language). There is a strong administrative effort to have a teaching staff that reflects the ethnic majority. Tev's first- and second-grade teachers use extensive multi-ethnic literature and talk with students about cultures. Nonetheless, there are racial slurs on the walls of the school, and a teacher's assistant gave advice to a child about "not letting a White girl beat you." In a culture circles activity, Tev did not draw a circle representing his Cambodian heritage.

Tev's second-grade teacher describes him as "not really below grade level. He's not one of my best. But he's really not below grade level. He tries. I can read enough of his stories that he's sounding stuff out."

In listing him as an example of an unsuccessful student, his first-grade teacher wrote that he was "extremely shy" and that she had been unsuccessful "in bringing him outward through activities" in class. She explained that he loves math and is very good in it, but that "he does not communicate." She added that his first-grade mentor "had been frustrated because she cannot get him to talk." During the Christmas concert at school, Tev knew none of the words to the songs, but stood on stage smiling broadly and moved his lips to the rhythm of the music.

In school, Tev seems bothered by little and bothers no one. He is said by his teachers to be "non-verbal" and "a very quiet child." At lunch, he eats little and talks little, but on the playground, he plays with a variety of the boys.

During a typical second-grade morning of drill-based work with writing and math, he is meticulous with the math ("making sure that his double digit numbers fit neatly under each other with long, straight black lines under them, and proper spacing for the total") and generally ahead of his classmates. The observer notes that when he copies sentences from the board, "his letters are neat and meticulous as his numbers, but his sentences come slowly and he often lifts his eyes from the paper to the board to copy the correct letters."

There are two places at school when Tev's voice is released—with his ESL teacher and with his mentor. Tev leaves and re-enters his regular classroom with virtually no notice from anyone. In his ESL room, the teacher suggests, "He gets his work done, raises his hand, and verbalizes more than in his regular class." The observer concurred. In a picture-based exercise designed to help ESL students develop categories and contents of the categories, Tev named nearly all the objects accurately.

Not surprisingly, Tev viewed his first-grade mentor with an "Uh-oh, here she comes again" attitude. His second-grade mentor, however, is a freeing and affirming influence for Tev. When the graduate student in economics arrives, Tev is demonstrably excited. The mentor feels Tev "is not as timid as everyone else has characterized . . . . We had an automatic bond, so that's why I say, you know, he feels comfortable around me, and I feel comfortable around him. So it wasn't where we had to sit there and pull teeth, you know." This mentor sees Tev's achievement as being "advanced." "I was surprised with his reading. He could read really well, and his math was pretty good too . . . . He's just a quick learner." In this setting, when he works with other students, he "talks and giggles." The mentor also sees Tev's strengths with the computer. "He knows how to go from menu to menu, how to use the mouse, even using the number pad and to click from one story to another."

Both Tev's first- and second-grade teachers say Project START has made them "more aware." His first-grade teacher says, "I feel like I was doing a lot of things already, but I wasn't really thinking about them or writing them down in categories like linguistic and logical-math, and it's made me do a lot of research too I'm constantly searching for new activities to do for spatial . . . . So I order a lot of manipulatives and things for building and gears and things like that to really tap into the spatial intelligence."

She has also worked to do some concept-based instruction—generally apart from her centers that remain skill-based. She sees Project START lessons as apart from other teaching, but says, "With my Project START lessons, I just think it was more appealing to them, and they weren't tuning me out so much because they were doing (work) in a way that they liked to do it. It's hard. It's a lot of planning. But it's definitely beneficial . . . . It really has to be thought out in advance." She also says that this year, there's not a lot of time for spatial activities—one of Tev's strengths.

The principal of the school is a strong supporter of Project START who reads, studies, and attends conferences related to multiple intelligences and multiculturalism. "We as a society need to do a better job with understanding other cultures," she says.

Tev seems to be "coming out" of his shyness a bit, according to his second-grade teacher. He doesn't appear to be unhappy about any of his classroom activities, but with the exception of time spent with his mentor, he doesn't seem to be excited about them either.

Denisha

Denisha is an African American third-grader who qualifies for free lunch. She is one of two children of a single mother, but has close connections with an extended family. Denisha's mother who has had some college education lives with her two children in low income housing. Denisha saw her uncle shot and killed in front of her house when she was very young. This event seems to have shaped her life in palpable ways.

At night, she fears the darkness, and calls to her mother to sing to her so she won't be afraid. She likes being in control of events, which makes people (including her mother and teacher) talk about her capacity to be bossy. She has a strong sense of justice and a need to protect people to whom she feels injustice is being done. "One day this boy had to stay inside from recess because he had one red check (for misbehavior) and I had one red check (but I could go outside). And that wasn't fair. So I said I'm gonna stay inside because she wouldn't let him go outside." Denisha wants to be a lawyer when she grows up so she "can protect people." Her mother worries that she will be judged by her very dark skin or because she lives in the projects. "I want her to be judged by what she knows." She worries that Denisha is not getting as much support at school as she should be getting. She has attended all of the Family Outreach activities and says they have helped her understand her daughter better. The Family Outreach Coordinator (also a fourth-grade teacher) for the school, however, is resentful of her Project START duties, saying she feels as though she's been "made to do someone else's work for them," and that she can't plan too many family outreach activities because of the busy schedules of parents. Nonetheless, Denisha seems to admire her and says she wants her for her fourth-grade teacher. The coordinator is "amazed" by this, and in her interviews does not seem to recognize that Denisha may identify with her, at least in part, because she is one of the few African American teachers in the school.

Denisha's mother keeps in touch with the teachers to encourage them to "push" Denisha in school. Denisha's current teacher also believes that this mother has very high goals for her child. They have had extended telephone conversations on a couple of occasions. "She expects (Denisha) to be doing her best and has always been very supportive if I needed that."

Mother and teacher also agree that Denisha is a compulsive talker. "She talks too much," says her mother. "(When she was little) the only way we could keep her quiet in church was to give her pencil and paper, and that's how she learned to write. Started at the age of one. She's been writing her alphabet ever since." Her third-grade teacher notes, "She will tell you what she thinks, and she means it. If she hurts your feelings, she'll say she's sorry, but that's what she meant. If she's not happy with someone, she's a little more difficult to manage. But she's forthright in any case."

Denisha is also sensitive. Her mother explains, "She's like an angel. She senses my problems and my fear, my hate, or whatever. She like, 'Momma, are you all right?' I'm like, 'Yeah, I'm fine,' you know, and it irritates me sometimes because it scares me. I'm like, 'Oh God, this child. I don't get her. I don't understand her, and every day she's like a new challenge.'" Her teacher reports that if she gets caught misbehaving, "that really upsets her." The class has a procedure through which any student can place a problem on the class roundtable agenda. The group meets weekly to discuss problems the students cannot resolve on their own. "When a problem involving her is submitted, "that upsets her. She doesn't want that."

In Denisha's third-grade class, the observer reports, "more time and energy are spent disciplining and threatening the kids into line than actually teaching and guiding them." The atmosphere is "tight" for this child who seems to need a looser fit. She seems to daydream a lot, although her teacher realizes she knows the answers even when she appears detached. At one point, she leans back and says to no one in particular, "I was just thinking that I've been here 40 days. I wonder what it will feel like when I have been here for 60 days."

During a science lesson, Denisha tries to engineer a group project, but her two teammates become weary of her suggestions and begin a glue gun war. When time is called, her group is the only one that has not finished the work. Only her simple drawing adorns the collage. She tells her mother, "I don't have any friends at school, don't nobody want to play with me." Her mother asks, "Well, are you trying to boss all the games and be controlling?" Denisha responds, "No. I just be me." The mother explains, "I'm like, 'Oh, God! (chuckle) 'cause I know her, and she'll tell you. If there's something wrong, she's gonna let you know . . . . She's bossy. I'd say she's a bossy little kid."

In second-grade, her teacher responded to Denisha's "attitude problem and impatience" by assigning her to an aide for special work. The observer concluded that the extra attention and focus may have been a positive for Denisha, who, this year, "appears to be floundering alone in an atmosphere where discipline and order are the cornerstone of the class, even at the expense of the lesson. Accompanying this dynamic is the feeling that they never quite get down to the lesson." Denisha tells her mother she is bored in class. That Denisha is an excellent writer is the only way her third-grade teacher speaks of her in a purely positive way.

Denisha, her mother, and second-grade teacher talk about how much she enjoyed her mentor in grade two. Her mother recalls, "I know she loved (the mentor). They were so much alike. Their birthdays were a day apart. They liked the same subjects, and after her visit, (Denisha) would talk about her for the rest of the week." Denisha herself says, "I liked leaving my classroom and working on projects with her." During third-grade, mentorships were late beginning. By December, she had only met with her new mentor once.

Denisha's second-grade teacher appeared "extremely committed to curriculum recommendations from START." She explained in detail how she used multiple intelligences as a way to organize her planning, how it brought out the creativity in her students, and re-energized her as a teacher. She accumulated a private library of books with a multicultural emphasis, which she encouraged students to read in the room or to check out.

Her third-grade teacher reported that, "Project START is probably more in line with what I personally would prefer doing the majority of the time. But the (district mandated) testing we are put through . . . sort of inhibits me from doing as much of it as I would like . . . so I'm just putting my toes in the water, so to speak, and seeing what fits best for me right now."

While Denisha talks of boredom in her regular classroom, she tells her mother she loves going to the special class for gifted learners for which she was recently identified. The teacher of this class, however, sees her as having potential more than performance. Commenting on a field trip related to architecture, the teacher noted that Denisha was visibly disengaged throughout the visit and did not function as a part of the group. He commented that if she lives in an environment where she simply appreciates shelter and safety, "this type of subject matter may seem superfluous and irrelevant to her." The teacher did not suggest making modifications in the subject matter to address the issue of relevance for Denisha.

Denisha does not seem to understand the concept of varied intelligences, but rather talks about strengths in discrete subjects or areas. "Most time I'm good in math, reading, social studies, and basketball . . . singing, 'cause I used to sing on children's choir, or praying."

Her mother says her strengths probably lie in spelling or math, "because she loves to spell and she loves numbers as well. She'll say, 'Well, I can't do this,' and I'm like, 'Okay, explain it to me, what you have to do.' And by the time she finishes explaining it to me, I'm lost and she knows how to do it . . . . I think in one way she's bright, but (Denisha) herself, she's like an angel."

Impact of Project START on Students and Parents

Qualitative research does not claim generalizability. Therefore, it is not the intent of this discussion to generalize to the entire Project START population or to other populations. Nonetheless, the eight case studies (four cases found in this report in detail; the other four were also used in analyses of themes) individually and collectively point to impacts on parents and students participating in Project START, as well as hinting at some cautions and considerations that may be useful to other program developers and implementers in similar settings.

1. Without exception, the eight students had at east one parent who staunchly and consistently supported them and were proud of them. Said one mother, "There was one extra sparkling star out there that shined on my baby and made her go. She reached up there and took her star. It was hers to have, she took it, she kept it, and now it's hers for life, and it's not going anywhere." Despite barriers of language, poverty, fear, and distance (which were factors in the lives of all eight families), the parents believed in their children and wanted to be a part of making their lives better. Such parents may be especially wise sources of investment for schools with meager resources. They did not always know how to parent well, but they wanted to know. They did not always know how to help their children with schooling, but they wanted to know. They were appreciative of even modest efforts, were active learners themselves, and were ready partners with school personnel. Without exception—whether their children were judged to be "successful" or "unsuccessful" in Project START—they defied the stereotype of uninvolved and unreachable parents.

2. Without exception, the eight START students and their parents felt the program made a positive and important difference to the student involved and/or to the family. While many of the students and parents were unclear about the meaning (or existence) of multiple intelligences, they were quite clear that the school was sending them a message that the child was important, smart, and worthy of attention. The message was not lost on parental self-esteem either. As one parent noted, "The school thinks my kid is smart. That must mean I'm okay, too—that I'm doing something right." For many of the parents, a formal communication from the school indicating that a child had great potential that the school wanted proactively to develop was a first. For one parent, it was frightening. She didn't show her child the letter because she feared it would turn out to be a mistake.

3. For the teachers, the concept of multiple intelligences served as a "boundary breaking" metaphor. That is, talking in specific and concrete terms about a range of intelligences affirmed for the teacher—at least intellectually—the possibilities of children whose capacities seemed "unorthodox" or beyond the traditional schoolhouse conception of "smart." The teachers seemed to accept, with relative ease, the notion that Project START students had potential which they could build upon, making the teachers believers rather than doubters where these children were concerned. Early in the first year of Project START, a teacher reflected, "Once I would have seen these children different with problems. Now I see them as different with potential." That attitude, albeit subtle, may well have immense power to communicate positively to students (and parents) whose lives are cluttered with problems. This attitude shift was clearly a contributor to improved school performance and/or attitudes about school among these students.

4. The team approach—or at least multi-avenue approach—used by Project START turned out to be powerful in an unanticipated way. The Project assumption that teachers can have a positive impact on young learners from the classroom, mentors from the community, and families can be strengthened through family outreach appears sound as reflected in these eight students. However, in no instance did all components work well for a given child. Sometimes a teacher was a strong advocate for the child, the mentor a real encouragement, yet family outreach worked less well. In other instances, the mentor and family outreach components were effective, but one or more of the child's teachers appeared less effective with the Project START learner. Having involvement with several key adults and program components means that if one element is weak, the capacity of the program to have a positive impact is not extinguished. For all eight of these students and their families, when a portion of Project START was "broken," other elements still functioned, having a net positive effect for children and families.

What Worked for the Project START Students

Mentorships

For the eight case study students, the most overt enthusiasm—from students, parents, and even teachers—was for the mentorship component of Project START. While the mentorship program was relatively modest (e.g., sometimes starting later in the year, typically meeting no more than once a week, meeting generally in the school rather than outside school, inconsistent linkage with classroom activities), it was concrete, visible, and positive. Students often indicated that they liked being singled out to leave the class for something peers would like to have the chance to do. It seemed to be a signal read by the mentee and other students that participants were important in the eyes of the school. For many, if not most, of these students, this may have been the first such school experience they had had.

In addition, the mentors themselves were often dedicated to and effective with their mentees. Tev's mentor gave him a voice. Belinda (a case study student not described above) was paired for two years with a local television meteorologist who was devoted to her young charge. Hungry for attention, Belinda responded eagerly. On Tuesdays (mentor days), unlike other days, she was never late or absent. She dressed up and wore a special perfume. She worked hard in her classroom on those days to ensure that her group got no "misconduct points." Belinda's second-grade teacher, the mentor coordinator, and Belinda's mother all attribute a positive transformation in her behavior and interest in school to "the TV weather lady." "She is able to zone in on (Belinda's) interests. There has been a big change in her attendance and tardiness." Prior to the mentor relationship, Belinda consistently refused to take part in physical education. The mentor encouraged her. "Now (Belinda) is walking the mile, running it, skipping around it, hopping around it, and just a lot of different things. She's developed coordination, which she should have developed a long time ago." Her mother also explains that Belinda used to become frustrated when she could not do things quickly and flawlessly. "(The mentor) taught her, it's not 'I'm gonna die if it's not done.' I can take my time and do this. (Belinda) was rushing through everything and nothing was getting done right. Then she slowed down and took her time." She says the mentor has also taught her daughter, "the importance of when you tell somebody you're gonna be there, be there. Okay. Don't come up with excuses. Just be there." Belinda's mother also believes her daughter is better in math because of strategies taught her by the meteorologist.

The mentorship for Belinda also tapped into her curiosity. "She loves to learn things she doesn't know, that she never even knew existed. (Belinda) didn't even know what a meteorologist was, and now she's all about learning about what a meteorologist is, what is the weather gonna be, how does (the mentor) get her information? What tools do they use. How do they know this is right?"

Belinda is allowed to stay up at night to watch her mentor on the late news so she can learn more about the profession. The power of knowledge has translated into family power as well. "When a seven-year-old can sit there and tell a twelve-year-old something she doesn't know, tell me they don't feel special. Very!"

Family Outreach

While the quality and consistency of the Family Outreach program varied from school to school, and while different parents had differing participation records, all of the case study students' parents knew about, appreciated, and attended at least some of the family outreach functions. In every instance, the parents were open to learning what was shared with them at the meetings. They also liked it when their children performed during those events. "The children really shine. It makes me feel good."

For most of the parents, family outreach activities were the only school activities they attended. Transportation provided by the school to the functions was a help. The lure of pizza suppers and door prizes also seemed helpful, but the parents mentioned these elements only casually. What they talked about was the chance family outreach provided to take part in activities with their children, and the chance it gave them to be more effective parents. Belinda's mother, for example, suggested that through family outreach, she learned about agencies she could call when she needed help for her children. "They (also) taught me it's okay to get angry. It doesn't make me a bad person, doesn't make me an abnormal parent. Makes me a very normal parent. And then (they taught me) to release my anger in a positive way. (They taught me) how to sit down and talk to the kids, put myself on their level and to talk to them like, 'We can work this out together.' They taught me it's okay to say you're sorry to your children, that is doesn't make you less of a parent, doesn't make you littler in their eyes . . . . Also how to deal with peer pressure, how to use positive punishment and discipline . . . . If it hadn't been for Project START, I just don't know (what would have happened to us). A specific strategy shared through family outreach that this single mother of four, who works two jobs, now uses to lessen tensions at home is an "attitude box." When anyone walks in the house with a problem, they [sic] write it on paper, drop it in a box, and literally leave it at the door. Later, they sit down together, look at the problems in the box, and find solutions together. The family also uses "time out" as a strategy that they learned through family outreach to calm tensions. When someone is frustrated and negative tensions are mounting, a family member will suggest that person take a time out so he/she can regroup. This mom talks about "things we've gathered and learned and nestled for hibernation out of Project START . . . (things) nobody can take away from us."

Classroom Modifications

Instructional adaptations advocated by Project START are the least consistently positive element of the project for the eight case study students. This should not be surprising in that teachers were being asked to change attitudes and practices in ways that caused major dissonance for many of them, with modest staff development, and in a shorter period of time than change literature suggests is feasible. Further, teachers in this district were under heavy mandates to raise student test scores on skills-based competency tests. This particular pressure made it even more difficult for many Project START teachers to depart from a drill-and-skill approach to teaching in favor of a more concept-based, sense-making approach advocated by Project START.

Some of the 16 teachers (eight children times two years) whose students were studied were eager subscribers to Project START instruction. One teacher found that the project provided her with models and strategies to create the student-centered classroom she had always wanted but didn't know how to achieve. One day when a researcher was in her classroom and students were working intently in several groups around the room on tasks of their own design, she talked about a child who had brought in materials from home as a result of earlier study and then organized peers to work with him on the materials. She also described a boy reading contentedly in a rocking chair, and then commented, "I've always dreamed of having a classroom that looks like this, but until this year, I never knew how to make it happen. I am so happy with my teaching now." This veteran of 15 years made major changes in her instructional routines, using not only multiple intelligences but also concept-based instruction in her planning and teaching. Several of the other teachers of case study students also expended concerted and consistent effort to plan lessons that were accessible via varied intelligence strengths, used learning centers more consistently, and learned to offer students more choices in their learning. Principals, researchers, and school resource teachers (e.g., support teachers for students identified as gifted) all commented repeatedly that Project START classrooms in general were more active, student-centered, and engaging than non-START classrooms. One support teacher conjectured that participation in Project START gave teachers "justification for trying new things . . . for loosening up their teaching style."

On the other hand, many of the 16 "case study" teachers made few consistently observable adaptations in regard to Project START instruction. Some came to understand how they might change and to develop a sense that making adaptations could benefit their students, but only "got their toes in the water" when in came to translating understandings into action. Others remained resistant to even the idea of change, suggesting that such changes took too much time, were counter to the district's testing mandates, or that they did not understand what the project was about. (This latter assertion was most defensible for third-grade teachers who only entered the project during its last year, and who did receive considerably less guidance and coaching in the instructional components of Project START than did first- and second-grade teachers.)

For most of the teachers whose students were in the case study group, "language immersion" remained "letting the kids talk a lot," manipulatives remained the property of math instruction, and multiculturalism was largely reflected in diversification of literature and study of holidays. These practices often had less to do with teacher rejection of principles than lack of time to become comfortable and conversant with extensive changes indicated by a student-centered, multiple intelligence, concept-based, multicultural, language-immersion, manipulative-oriented classroom.

Differences and Similarities in "Successful" vs. "Unsuccessful" Project START Students

These students make a compelling case for the semantic complexities of a construct like "success." Most of the parents appeared to judge their children to be successful—they were "smart" and "hanging on," even in the face of stunning difficulty. Two students judged successful by their teaches at the end of year one in Project START were judged to be "talented in many areas." Both were female, compliant, and unobtrusive. The other two nominated as successful (and selected for the case studies) were listed by their teachers because of progress in self-esteem and behavior. The male was said to engage in "less fighting" at the end of the year than at the beginning, and to have "checked his attitude." The female was noted to have "a better attitude" and "more complete work."

Nominating comments on students deemed by their teachers to be "unsuccessful," and selected for the case studies, were more mixed. Tev was said to be uncommunicative (no suggestion in the nominating comments or interviews that the language barrier might be the cause). A male was said to be unsuccessful because he talked a lot, but in unproductive ways, and because he had behavior problems that had not been able to be channeled in positive ways. A female (Denisha) was said to be unsuccessful because she had "become more demanding of attention," seemed "to feel she deserved special consideration," and was "not putting forth her best effort." The fourth "unsuccessful" student selected for the case study was said to have "demonstrated minimal talent."

Test scores (Iowa Tests of Basic Skills), not likely to best reflect growth in non-traditional student populations, did affirm teacher conceptions of success (that is, showed marked growth during the course of year one of the project), in three instances of non-success (that is, showed little growth, no growth, or negative growth during year one of the project). In two cases, the teacher judged a child unsuccessful when the ITBS indicated marked growth. In two instances, the teacher judged a child successful when the ITBS indicated little, no, or negative growth.

In general, among these students and teachers, a child was more likely to be judged successful by a teacher if he/she (a) demonstrated outstanding ability in "traditional" areas (e.g., writing, reading, spelling, math) and/or (b) was a "low maintenance" student (e.g., was compliant and had few or no behavior problems). Also in general among these students, a child was likely to be judged unsuccessful if he/she (a) demonstrated talent predominately in non-traditional areas (e.g., spatial, interpersonal) and/or (b) was a "high maintenance" student (e.g., required considerable intervention from the teacher to deal with complications caused by language or behavior). The former pattern suggests that, not surprisingly, teachers were unable to rapidly change their conceptions of intelligence. The latter pattern suggests that in classrooms where teachers must deal with so many students, problems, and decisions, a child who challenges the system becomes a stressor and simply doesn't measure up to the prevailing norms for success.

It is important to note, however, that in at least three instances, students deemed to be successful or unsuccessful by their year one teachers were viewed as opposite by their year two teachers. In these instances, case study observers felt the reversal was more a product of teacher differences than changes in the children.

In actuality, the case studies indicate that all eight of the Project START students "succeeded" in noteworthy ways: increased attendance, enthusiasm for schoolwork, and growing self-esteem, greater language skills, increased test scores, and identification for the districtwide program for students identified as academically gifted. All of the eight students succeeded in several of these categories during the three-year duration of the project.

CHAPTER 5: General Discussions and Implications

One of the most important challenges facing gifted programs and the field of education in general is the development of the full potential of all children in our schools. The responsibility for these efforts will ultimately rest with classroom teachers and their skills in identifying talents and in implementing an instructional program that maximizes the development of all children's potential. For many yeas, writers in the field of gifted education have called for broadened conceptions of giftedness, for alternative means of identifying talent, and for new programs to enhance talent in order to achieve that goal (Grinder, 1985).

Minority students, particularly those from low-income backgrounds, are typically among those whose talents are not recognized nor fully developed as evidenced by the under-representation of these students in programs for gifted learners (Baldwin, 1994; Frasier & Passow, 1994; U.S. Department of Education, 1993). This is the case for an array of reasons, including: limited and limiting definitions of giftedness, lack of culturally sensitive means of assessing potential, inadequate preparation of teachers in issues and practices related to an increasingly multicultural school population, and lack of awareness among teachers about ways in which potential in culturally diverse populations may be manifest (Ford, 1994b; Frasier, Garcia, & Passow, 1995). Further, once identified for services provided for gifted learners, students from minority cultures may fare poorly in programs that overlook cultural differences in learning (Banks & Banks, 1993; Tharp, 1989). While there is much diversity within any single culture or economic group, and it is thus important not to overgeneralize about individuals within such groups, researchers have found patterns of learning in varied cultural and economic groups that illuminate the reasons many minority students are less successful. Their experience may mitigate against meeting the expectations for learning in traditional classroom environments (Banks, 1993).

One paradigm that has been proposed is the adoption of new conceptions of intelligence and intelligent behavior (Gardner, 1983; Renzulli, 1978; Sternberg, 1985) as the basis for teacher recognition and identification of highly able students and the provision of alternative, more appropriate instruction. Gardner's conception of intelligences, ranging from the more traditionally recognized verbal-linguistic ability and mathematical-logical ability to spatial, inter- and intrapersonal, and musical ability, have been translated into assessment protocols and instructional programs (Gardner, 1991b). Despite widespread acceptance of this model and recommendations for its use, few studies have been undertaken to provide insight into the efficacy of alternative identification methods using this model or about the experiences of culturally diverse youngsters identified or served as gifted using the model.

The Study

Project START (Support to Affirm Rising Talent) was a three-year collaborative research effort undertaken jointly by the University of Virginia site of The National Research Center on the Gifted and Talented (NRC/GT) and the public schools of Charlotte-Mecklenburg (North Carolina) (CMS). While approximately 40% of the Charlotte-Mecklenburg students were minority, fewer than 10% of identified gifted learners were minority. Approximately 69% of the African American elementary school population of the district were from low income backgrounds at the time the study began.

The researchers at the NRC/GT and the CMS school personnel agreed to develop a program based on Gardner's (1983) multiple intelligences theory. The specific purposes of the study were to: (a) develop identification procedures based upon Howard Gardner's theory, (b) identify high-potential primary age students from culturally diverse and/or low economic backgrounds through use of Gardner's framework, (c) investigate the reliability and validity of the identification procedures, and (d) test the efficacy of specific interventions on achievement and attitudes about school and self of identified students.

The students were identified using protocols developed by CMS. Identified students were assigned to one of three conditions: (a) an experimental condition that consisted of modification of classroom activities and a family outreach program, (b) an experimental condition that consisted of modification of classroom activities, a family outreach program, and a mentorship, or (c) a control group. Assignments to the treatment groups were made according to the school the child was attending. Schools had been randomly assigned the condition. Control group students were in all schools, but their teachers were not trained in the intervention procedures, the students did not have a mentor, and their families were not invited to participate in the family outreach program.

The Implications of the Findings

The Psychometric Properties of the Assessment Tools

Reliability evidence suggests that the subscales (logical-mathematical, linguistic, spatial, interpersonal) are internally consistent. The factor analysis confirmed the presence of the linguistic and logical-mathematical subscales, but the presence of the two remaining subscales could not be confirmed. The combination of linguistic and interpersonal intelligence activities on the first factor is not surprising, since interpersonal communication contains a major verbal component. Analysis of the Interpersonal Checklist (Udall & Passe, 1993) reveals that many of the items on the checklist can be interpreted with an emphasis on verbal-linguistic talent (e.g., "Acts as peacemaker," "Shows humor in interactions," "Is listened to and sought out by other children"). The split of spatial activities between the third and fourth factors is more puzzling. Activities 2 (Pablo) and 4 (Tangrams) are very similar, so their loadings on the fourth factor were expected, but their separation from the other spatial activities is not easily explained, except for the strong geometric orientation of these two activities. Since the fourth factor accounts for little explained variance (when a three-factor model is used, the third and fourth factors collapse), the split of spatial activities may have little practical significance.

No meaningful gender differences were found on the assessments, but the relatively high ratings of Asian American students are cause for concern. Teacher subjectivity may be influencing the assessments, or the use of performance assessments to avoid ethnic bias on standardized tests may simply be misguided. Significant differences in ratings among schools exist (although they are associated with small effect sizes), which supports previous research on the inconsistency associated with performance-based assessments (e.g., Aschbacher, 1991; Haertel, 1994).

Analysis of the multitrait-multimethod matrix yielded limited evidence supporting convergent and discriminant validity of the assessments. A possible explanation for the relatively large correlations among the mathematical-logical assessments (tasks and checklists) and ITBS subtests may be that the activities and checklist associated with the math-logical subscale are the most objectively scored of those in the MI-based alternative assessment battery.

Establishing evidence of concurrent validity of new assessment tools based on alternative assessments of intelligence presents a difficult challenge to test developers and educators who wish to use the assessments. The construct validity issue becomes immediately apparent as we examine our data. On the one hand, the definitions of logical-mathematical intelligence and linguistic intelligence lead us to expect that ability in these areas would correlate with high scores on achievement tests that assess outcomes relating to high ability in those areas. On the other hand, the writings of Gardner (1984) are clearly critical of traditional assessment tools as being too narrowly conceived to capture the richness of aptitude and performance. If this is true, then how can the validity of MI-based assessments be established? Is MI theory essentially unable to be proven? Gardner and Hatch (1990) address this issue directly:

> Some critics have suggested that MI theory cannot be disconfirmed. . . . If future assessments do not reveal strengths and weaknesses within a population, if performances on different activities prove to be systematically correlated, and if constructs (and instruments) like the IQ explain the preponderance of the variance on activities configured to tap specific intelligences, then MI theory will have to be revamped. (p. 8)

We strongly believe that, for legal, educational, and ethical reasons, performance assessments used for high-stakes purposes such as identifying potentially talented students need to be reliable, valid, appropriately normed, and equally fair to students regardless of gender, ethnicity, and SES. This belief is voiced by others who are concerned with the proper use of alternative assessment techniques (O'Neil, 1992). Miller and Legg (1993) believe that "when the stakes for assessment scores are high, the traditional notion of valid and reliable interpretation of scores remains critical" (p. 10), and Messick (1994) argued that:

> [S]uch basic assessment issues as validity, reliability, comparability, and fairness need to be uniformly addressed for all assessments because they are not just measurement principles, they are social values that have meaning and force outside of measurement wherever evaluative judgment and decisions are made. (p. 13)

The revised assessment battery is reliable for group purposes as well as research purposes, although the entire collection of activities does not appear to be sufficiently reliable for individual administration and interpretation. With increased training and streamlined logistics, however, the reliability of the assessments might be increased. The greater concern involves the validity of the assessment battery.

The weak relationship among the ITBS subtest scores and alternative assessments in the first correlational multitrait-multimethod analysis suggests that the battery is characterized by low convergent validity. Another explanation is that the tests are assessing completely different aspects of mathematical and linguistic intelligence or aptitude than the MI-based assessment. Egan and Gardner (1992) described this difference as "knowing" (i.e., the ITBS) versus "understanding" (the performance assessments). The moderate correlations among the performance activities and teacher/observer checklists and rating scales could be used to support the position that evidence of moderate convergent and discriminant validity exists only within the battery and not across different assessment methods. This phenomenon may explain the lack of significant, quantitative results of some MI-based projects with respect to student achievement. That is, performance, for example, on the ITBS, is depressed because the MI training focuses on different concepts.

The results of the confirmatory factor analysis must be interpreted with caution. Although considerable support exists for a refutation of Gardner's model, the same evidence can be viewed as partially supporting MI theory. In the best fitting model generated by the factor analysis, the four highly correlated MI-based constructs are further organized by a general, second-order factor. Thus, the best fitting model could be interpreted as supporting a "g"-based theory of intelligence. Conversely, the hierarchical model suggested in the analysis may be used as evidence that in the process of rating students the observers and teachers rate and score students making some minimal differentiation across the four MI constructs, but that the differentiation they make is inhibited by a tendency to be influenced by the method of scoring (ratings on the scales) and by limitations imposed by their own biases about intellectual ability. That is, they may tend to be influenced in rating any child's abilities in any dimension by their perceptions of the child's linguistic and/or mathematical/logical abilities. Further, the wording of the scales may not provide the guidance necessary to over-ride that tendency.

The presence of relatively large ethnic differences is also cause for concern. In contrast to the previous administration of the battery, in which ethnic differences were small, the gap between the scores of Caucasian and African American students was much wider. This raises several questions about the assessment process: Are teachers more likely to fall back onto streotypic judgments of student behavior as they become more familiar with alternative assessment procedures? Does a point of diminishing return exist after which additional inservice training with alternative assessments has no appreciable effect? As the answers to these two questions have potentially severe financial, logistical, and educational implications, they should be investigated in the near future.

This study unequivocally proves one point: Testing MI or similar theories of multiple intelligences is difficult, and designing instruments to test MI theory or assess multiple intelligences in students involves significant problems that individuals investigating other alternative assessments or theories of intelligence may not encounter. As Eisner (1994) commented,

> While Gardner does point out—to his credit—that high-level work in all fields usually requires more than one form of intelligence, the fact that the vast majority of the population operates in the middle range may have something to say about the meaning of types of intelligence itself. It may be much more difficult empirically and conceptually to display the clarity of acute types in dealing with vast populations in which the mix is more balanced. (p. 559)

Doing so, in fact, may involve tremendous staffing and training resources that our schools cannot and will not provide. However, regardless of research results and eventual decisions about the correctness of MI theory, it still has value as a heuristic or framework for producing educational change. Gardner (1994) supports this application, stating that he believes MI theory "is less a set of hypotheses and predictions than it is an organized framework for configuring an ensemble of data about human cognition in different cultures" (p. 578).

Any theory or heuristic that allows teachers to take a more inclusive view of talent, intelligence, and ability has value in our schools and for our students. For the moment, MI theory may prove useful for this purpose, while applications to assessment clearly need to be the focus of future research efforts.

Quantitative Results

In summary, the results from the quantitative analysis of achievement data suggest that the intervention had no significant effect on the achievement of the treatment groups in any of the grades. All children in the study demonstrated growth and by the project end were, in most cases, on grade level.

However, for the students who were followed throughout the project one positive result was noticed: regardless of which treatment group the students (typically identified as at-risk) were in, they were on grade level by project end. Although no particular treatment can lay claim to this result, perhaps there was some "filtering out" to all teachers, including non-START teachers, of the project goals that consequently influenced these students' classrooms.

On attitudes toward learning, teachers, language arts, and mathematics, the students in all three groups in cohort 1 experienced a decrease in attitudes toward learning and teachers. In cohort 2, attitudes toward teachers and language arts increased significantly for all three treatment groups, but attitudes toward learning decreased.

In the area of self-concept, little consistency occurred across the two cohorts. However, in cohort 2 those students who were assigned a mentor showed consistent increases across the two time periods for all four areas of self-concept. This was the only pattern that emerged, although it was not statistically significant. This finding does suggest somewhat the positive effects that a mentoring program can have on at-risk students.

Qualitative Results

Teacher Change

Teacher instructional change appeared evident, but modest, as a result of Project START. At least three themes that may inform subsequent, research related to Project START ran through the teacher change data. One is more concrete and specific to projects attempting to use multiple intelligence theory related to talent development in students with varied intelligence strengths. The other two recurring themes are more abstract and genealizable to a broad range of projects involving teacher change.

This project indicates a clear need to further define ways in which multiple intelligence theory can be translated into classroom practice in ways which are understandable and attainable for a broad range of classroom teachers, without corrupting the theory itself. MI materials have been rapidly popularized and provide recipe-like lists of activities that purport to develop student intelligences. These materials seldom deal with issues of the nature of a discipline or the need for a coherent curriculum. They virtually never delve into what would constitute use of MI theory as genuine development of an intelligence or talent area compared to use of the theory as yet another view of student learning style. A sense of what beginning kinesthetic talent or spatial talent might look like compared with advanced kinesthetic or spatial talent, for example, is largely absent in most MI teacher materials, as is clear guidance on how strengths in one area might effectively be employed to help a child develop authentic strengths in an area of weakness. This project identified some of the gaps in thinking regarding translation of MI theory into classroom practice and began filling in a few of those gaps. Much more needs to be done before MI becomes a popular panacea. Currently, theory far outstrips articulated understanding about how that theory might be translated into defensible practice.

One of the more abstract themes evident throughout Project START relates to complex issues of site selection in educational research and what is ethical to expect of teachers in selected sites. Project START appeared well suited to the Charlotte-Mecklenburg Schools in that both shared a driving goal of boosting the achievement prospects of at-risk learners. In time, however, it became clear that the school district had charted a route for achieving the shared goal that was almost diametrically opposed to the route through which Project START sought to achieve the goal. A case can doubtless be made for either route toward the commonly desired end. While Project START researchers clearly had a preference for one of the avenues, the intent is not to argue for one means over the other as much as to raise the question of what might be done in selecting research sites (a) to ensure that not only principles but practices are shared between researchers and participating districts, and (b) to free teachers from simultaneously competing demands of two powerful groups of leaders. In one after-school workshop during year two, Project START teachers were given a variety of materials and modes of expression (modeling multiple intelligence practices) and asked to share their current sense of their work with the project. Without collaboration among the individuals and groups, the vast majority of expressions focused on the teachers' sense that they were being pulled in two directions at once—"pulled apart at the seams."

In this district, the task of reducing the competing demands on participating teachers was magnified by site-based decision-making, in which individual schools could select curricula, make instructional decisions, and impose duties as school leadership saw fit. Again, it is not the purpose of this reflection to argue for or against site-based management, but rather to note that research site selection is confounded when each site may operate in ways quite different from most other sites in a research project. Researchers should consider negotiations with potential collaborative school districts through which those districts agree to free teachers from competing mandates for the duration of the research, based on the belief that testing varied methods of achieving success within a given district has merit. In addition, teams of researchers should explore ethical responses to research requirements placing teachers in positions that cause them to have to choose whether to support the research project on which they have agreed to work (and in whose practices they may believe) or the school district that hires them, evaluates them, and controls their salaries.

A third somewhat abstract issue raised by the teacher change component of Project START, and potentially significant to related educational research involving teacher change, is the issue of project scope. Educational reformers broadly and convincingly argue that it is ineffective to continue to apply Band-Aids to gaping wounds—that is, to use a quick instructional fix that falls short of addressing endemic classroom problems. It also appears evident that making significant changes in educational practice involves teacher change of some breadth and depth. These tenets lead classroom researchers to design interventions broad in scope and potentially wide ranging in ramifications. Grant funding cycles are often briefer in duration than research indicates is necessary to make significant changes in teacher practice. The brevity of funding cycles is especially problematic when student outcomes on standardized instruments are sought as evidence of the efficacy of the intervention, and those outcomes are predicated upon teacher change. Three means of addressing the problem of project scope versus duration of funding seem evident—if difficult to achieve. First, balancing the value of outcomes measured by reliable and valid quantitative research with outcomes described with trustworthy and credible qualitative research offers two very useful perspectives on complex classroom settings. Second, developing funding options that encourage renewal of research grant funding for a second cycle in complex projects would seem more likely to result in useful long-term understandings of the impact of complex interventions on teacher and classroom change. Third, negotiations with school districts prior to agreements to collaborate in which the district and/or researchers agree to continue expending time and fiscal resources in applying the intervention and studying its impacts beyond a necessarily short funding cycle would also facilitate understanding of the long-term efficacy of the interventions, as well as supporting teachers as they continue in the arduous process of change.

Case Studies of Students

Early examination of the quantitative data indicated great variability among the students in degree of success in the program. Further, teachers commented frequently on degrees of success of identified students. The test data did not give us an explanation for these differences or insight into the connections between home and school that are important in any child's life, but may have particular bearing on lives of students for whom poverty and/or cultural minority status are complicating variables. It also did not give us an opportunity for insight into ways the program components, either singly or in combination, impact the lives of students.

Some Final Thoughts

Once again, while not claiming generalizability, the Project START cases suggest several themes for future researchers on related projects to consider.

1. Educators must develop images of high-potential, high-risk learners that extend well beyond the classroom. These students and their families have complex lives that—positively or negatively—profoundly affect relationships at school and engagement in learning. For example, Tev is seen by many at school as non-communicative and shy. Strange as it seems, little, if any, connection was articulated by school personnel between his family's lack of comfort with English and his own reticence. Similarly, little note seemed to be paid to his growth from ground-zero to his current developing fluency. Judged on an absolute standard, he becomes a non-entity. Judged on personal development, he is anything but unsuccessful.

   Likewise, Denisha is seen as combative by many adults in her school setting. While that may, in fact, be the case, dealing with her combativeness as a somewhat volitional behavior—instead of as the result of a terrible fear born of early violence in her life—might well result in adults who resent her behaviors and address negativity with negativity, rather than adults who understand her need for control and justice, and thus work with her in more positive ways.

   One mother, not minding her language, understood. "You know, they can lash out in anger because they're all pissed off because daddy got up and left years ago, and it's just affecting them." In spite of family outreach programs that were a plus, case study researchers received indication that school personnel were aware of and addressing the specific "loaded" circumstances of these children's lives.

   Merely developing pre-project case studies, such as the ones completed later in Project START for this report, could provide sensitive educators with powerful insights into learners' lives and help teachers craft environments that ameliorate rather than ignore or complicate already challenging circumstances. In any case, linking knowledge of the child at home to knowledge of the child at school is important. Certainly a rich interpretation of a multicultural curriculum would be enhanced by educators extending their understanding of the cultures from which their learners actually come.

2. In project design as it relates to teacher change powerful enough to predict substantial impact on the learning of high-risk, high-potential learners, at least two themes recurred among teachers of the case study children. First,

if such teacher change is to occur, substantial time and consistent coaching are required. That notable modification did occur in some Project START classrooms has been addressed earlier. However, many teachers were just beginning to enact change as the funding cycle ended. One teacher said as the project ended, "I think I'm beginning to catch on here, but now you're going away." District support beyond the grant funding in ways that continue to affirm and extend change begun during the grant cycle seems imperative in order to be fair to teachers who risk change and in order to avoid a prevalent sense of revolving door mandates that come and go with little impact other than disruption in teachers' lives. Further, examination of case study teachers suggests a need for on-going, on-site coaching of teachers toward desired change throughout the funding cycle. Staff development that is concentrated in a week in the summer and monthly after school meetings is likely to be more effective only in developing "knowledge about" than "knowledge how."

A second theme among teachers of case study students is the need to select sites for funding where district and/or school initiatives are in harmony with goals of the grant being funded. Project START goals included developing student-centered classrooms in which students have opportunity to construct meaning via high relevance content and activities, using non-traditional assessment mechanisms, and providing numerous avenues to learning so that children with strong talent in non-traditional areas recognize themselves and their potential through the curriculum. On the contrary, the district in this setting had a heavy emphasis on raising student test scores on traditional, skill-based tests—to the degree that teacher salaries are linked to ascending test scores, and as one teacher noted, "Principals disappear overnight if the test scores don't go up." Project START teachers found it confusing and difficult to embrace these competing approaches simultaneously. One noted, "What START is telling us to do makes sense to me. It feels right. But I'm afraid for my job if I try it and (students') scores fall." Not only were teachers compromised by being pulled in what they continued to perceive as opposite directions, but the issue of time to bring about teacher change was compounded due to the fact that the almost total emphasis on drill and skill found in primary classrooms was so far removed from a more concept- and meaning-based approach desired by Project START.

3. Several of the case study students were identified for and participated in the district's program for gifted learners, a major goal of Project START. Generally, support staff in the program for the gifted appeared enthusiastic about working with these students, interested in their progress, and aware of adjustment issues that face these students as they enter such a program. For example, several support teachers discussed pressure on African American males by peers to conform to group norms rather than be a part of academic success. Several also talked about disparity in Project START students' projects in comparison with projects completed by more affluent students who have a broader range of support mechanisms at home. Less evident was a specific plan for undergirding Project START student success in academically advanced programs and classes (e.g., peer support groups, after school work sessions for project work with transportation home provided, additional during-school time for project work with support teachers). There was also less evidence of differentiation of instruction in special classes for Project START learners (and others who might benefit from it) to address

> experiential gaps, learning styles, issues of content relevance, and so on. When programs such as Project START are created, at least in part to promote equity of identification for and participation in programs for students identified as gifted, it would seem important to emphasize success-building for those students who do participate in the services that open up to them as a result of the initiatives.

In Project START, as reflected in the eight case study students, time and support for teacher change was scarce, mentorships were only loosely linked with classrooms, and family outreach tended to make parents feel valued in special project activities, but seemed to stop short of involving parents in "mainstream" school events. Also, the notion of multiple intelligences served more as a metaphor for student possibilities than a robustly translated approach to extending student intelligences toward advanced production levels. Said in that way, Project START seems a less than complete effort at making a difference in the lives of Tev, Denisha, William, Charelle, and the other four case study students.

Studying these students in some depth and over several months, however, led to a different impression. Certainly given extended time and resources, more could be accomplished with each facet of the project and thus for the students. However, what appears evident in these children's lives—and those of their families—is that even modest affirmation and intervention such as the type inaugurated by Project START makes a real difference. When a teacher begins to think about a child in more positive than negative ways, when a classroom becomes more flexible, when a parent hears a message from school that a child is talented and worth special investment, when the doors to school seem open and inviting, when someone from outside the school comes and spends time with a child, important transformations occur. Teachers emphasize what is right in a child rather than what is wrong. Parents who feel oppressed by what does not work in their lives begin to see something that does work. Students who may face life "with their dukes up" because of the tensions that surround them find school a more inviting place and home a bit more hopeful. In these ways, in fact, there was, in varying measure, for all of the case study students, success that appeared directly linked to Project START initiatives. The process of bringing about school change is predicated on our ability to change teacher behavior in the classroom. In implementing Project START we confirmed findings of other researchers that the process is very complex and requires a long, serious, and consistent commitment on the part of the central and building level administrative staff.

On the one hand, our investigation of using the Gardner (1983) paradigm for identification of talent led us to conclude that the reliability and validity of such performance assessment tools are still quite tenuous and school personnel should be wary of making high stakes decisions such as identification for gifted programs based primarily on these measures. On the other hand, the paradigm did open teachers' minds to possibilities and promise in the children in their classrooms. While they often had a difficult struggle to see the differences between the use of the intelligences as a means of understanding learning capabilities rather than learning styles, they did come to see students in a new light and to treat them quite differently. This change in teacher attitudes may provide some explanation of the more positive quantitative changes noted in self-concept in the treatment groups.

Our findings also support the value of combining quantitative and qualitative data analysis. The quantitative data allowed us to examine overall impact on groups of children. Qualitative data, however, provided useful insights about teacher change and impacts of various project components on students and parents. These insights would not have been revealed by quantitative data alone.

As the staff of the project reflected on the results and implications of this study, we learned many lessons that are not necessarily those related to talent identification and development in particular, but rather to general principles of the change process in schools. First, we found that school administrators in desperate attempts to bring about high achievement may be institutionalizing practices that mitigate against the very change they are seeking. Our experiences with teachers demonstrated over and over that high stakes testing resulted in very anxious teachers who often responded with ineffectual behavior with little ability to judge its ineffectiveness. Rather than interpret negative outcomes as an indicator of need to try alternative solutions, they would increase the intensity of inappropriate behavior. For example, when drill and practice were not increasing test scores, they adopted more drill and practice. When low level review bored children who then disengaged from learning, they lowered the bar even more—creating even more simplistic, more boring activities that disengaged the children further.

Certainly the study highlights the complexity implicit in modifying schools and classrooms. It is difficult for all of us, teachers included, to understand the impact of culture and economic status on learning. Teaching is a heavily habitual endeavor. Changing any entrenched pattern of thought and behavior is demanding. Changing a network of such habits is confounding. Finally, teachers are often at the mercy of institutional policies and procedures. When those policies and procedures run counter to teacher beliefs, or when they inhibit teacher reflection upon beliefs and instructional practices, the likelihood of robust change is further diminished. Nonetheless, this study suggests that the lives of individual high potential, high risk students can be positively impacted even by modest and incremental changes in classroom practice and school climate.

References

Abeel, L. B., Callahan, C. M., & Hunsaker, S. L. (1994). *The use of published instruments in the identification of gifted students.* Washington, DC: National Association for Gifted Children.

Adams, M. L., & Feldman, D. H. (1993). Project Spectrum: A theory-based approach to early education. In R. Pasnak & M. L. Howe (Eds.), *Emerging themes in cognitive development. Volume II: Competencies* (pp. 53-76). New York: Springer-Verlag.

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology, 43*(5), 97-103.

Anderson, E. M., & Shannon, A. L. (1988). Toward a conceptualization of mentoring. *Journal of Teacher Education, 39*(1), 38-42.

Arkansas State Department of Education. (1991). *McRat report—multicultural reading and thinking.* Little Rock, AR: Arkansas State Department of Education. (ERIC Document Reproduction Service No. ED 337 766)

Arlin, M. (1976). *Manual for Arlin-Hills attitude surveys.* Jacksonville, IL: Psychologists and Educators.

Arnold, K. D., & Subotnik, R. F. (1995). Mentoring the gifted: A differentiated model. *Educational Horizons, 73*(3), 118-123.

Aschbacher, P. R. (1991). Performance assessment: State activity, interest, and concerns. *Applied Measurement in Education, 4*(4), 275-288.

Ascher, C. (1988). Improving the school-home connection for poor and minority urban students. *The Urban Review, 20*(2), 109-123.

Au-pei-Hu, K. (1980). Participation structures in a reading lesson with Hawaiian children: Analysis of a culturally appropriate instructional event. *Anthropology and Education Quarterly, 11*(2), 91-115.

Baer, J. (1994). Performance assessments of creativity: Do they have long-term stability? *Roeper Review, 17*(1), 7-11.

Baker, E. L., Aschbacher, P. R., Niemi, D., & Sato, E. (1992). *CRESST performance assessment models: Assessing content area explanations.* Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.

Baker, E. L., O'Neil, H. F., Jr., & Linn, R. L. (1993). Policy and validity prospects for performance-based assessment. *American Psychologist, 48*(2), 1210-1218.

Baldwin, A. Y. (1977). Tests can underpredict: A case study. *Phi Delta Kappan, 58*(8), 620-621.

Baldwin, A. Y. (1984). *The Baldwin identification matrix 2 for the identification of the gifted and talented: A handbook for its use.* New York: Trillium Press.

Baldwin, A. Y. (1987). I'm Black but look at me. I am also gifted. *Gifted Child Quarterly, 31*(4), 180-185.

Baldwin, A. Y. (1994). The seven plus story: Developing hidden talent among students in socioeconomically disadvantaged environments. *Gifted Child Quarterly, 38*(2), 80-84.

Banks, J. A. (1988). *Multiethnic education: Theory and practice* (2nd ed.). Boston: Allyn & Bacon.

Banks, J. A. (1991). *Teaching strategies for ethnic studies* (5th ed.). Boston: Allyn & Bacon.

Banks, J. A. (1993). Approaches to multicultural curriculum reform. In J. A. Banks & C. A. M. Banks (Eds.), *Multicultural education: Issues and perspectives* (2nd ed., pp. 195-214). Boston: Allyn & Bacon.

Banks, J. A., & Banks, C. A. (1993). *Multicultural education: Issues and perspectives.* Boston: Allyn & Bacon.

Barbour, N. (1992). Early childhood gifted education: A collaborative perspective. *Journal for the Education of the Gifted, 15*(2), 145-162.

Baruth, L. G., & Manning, M. L. (1992). *Multicultural education of children and adolescents.* Boston: Allyn & Bacon.

Beck, L. (1989). Mentorships: Benefits and effects on career development. *Gifted Child Quarterly, 33*(1), 22-28.

Bednar, A. K., Cunningham, D., Duffy, T. M., & Perry, J. D. (1991). Theory into practice: How do we link? In G. Anglin (Ed.), *Instructional technology: Past, present, and future* (pp. 69-92). Denver, CO: Libraries Unlimited.

Bennett, C. I. (1986). *Comprehensive multicultural education.* Boston: Allyn & Bacon.

Berger, S. L. (1990). *Mentor relationships and gifted learners.* ERIC Digest #E486. Reston, VA: CEC. (ERIC Document Reproduction Service No. ED 321 491)

Berman, P., & McLaughlin, M. W. (1979). *An exploratory study of school district adaptation.* Sacramento, CA: Rand. (ERIC Document Reproduction Service No. ED 175 089)

Bloom, B. S. (1985). *Developing talent in young people.* New York: Ballantine.

Blum, D. J., & Jones, L. A. (1993). Academic growth group and mentoring program for potential dropouts. *The School Counselor, 40*(3), 207-217.

Borland, J., & Wright, L. (1994). Identifying young, potentially gifted, economically disadvantaged students. *Gifted Child Quarterly, 38*(4), 164-171.

Bruner, J. S. (1986). *Actual minds, possible worlds.* New York: W. W. Norton & Company.

Burger, S. E., & Burger, D. L. (1994). Determining the validity of performance-based assessment. *Educational Measurement: Issues and Practices, 13*(1), 9-15.

Callahan, C. M., Tomlinson, C. A., & Pizzat, P. (1994). *Contexts for promise: Noteworthy practices and innovations in the identification of gifted students.* Charlottesville, VA: University of Virginia, The National Research Center on the Gifted and Talented.

Campbell, D. T., & Fiske, D. W. (1967). Convergent and discriminant validation by the multitrait-multimethod matrix. In D. N. Jackson & S. Messick (Eds.), *Problems in human assessment* (pp. 124-132). New York: McGraw-Hill.

Candell, G. L., & Ercikan, K. (1994). On the generalizability of school-level performance assessment scores. *International Journal of Educational Research, 21*(3), 267-278.

Clark, B. (1997). *Growing up gifted* (5th ed.). Upper Saddle River, NJ: Merrill.

Clasen, D. R., Middleton, J. A., & Connell, T. J. (1994). Assessing artistic and problem solving performance in minority and non-minority students using a nontraditional multidimensional approach. *Gifted Child Quarterly, 38*(1), 27-32.

Clifford, J. A., Runions, T., & Smyth, E. (1988). Educating the gifted adolescent as co-learner: The learning enrichment service (LES) and the mentor academy program (MAP). *Gifted Education International, 5*(3), 160-169.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum.

Collins, A. (1993). Performance-based assessment of biology teachers: Promises and pitfalls. *Journal of Research in Science Teaching, 30*(9), 1103-1120.

Collins, E. G. C., & Scott, P. (Eds.). (1978). Everyone who makes it has a mentor. *Harvard Business Review, 56,* 89-101.

Comer, R. (1989). A mentorship program for gifted students. *The School Counselor, 36*(3), 224-228.

Comrey, A. L. (1973). *A first course in factor analysis.* New York: Academic Press.

Cox, J., Daniels, N., & Boston, B. (1985). *Educating able learners: Programs and promising practices.* Austin, TX: University of Texas Press.

Csikszentmahalyi, M., Rathunde, K., & Whalen, S. (1993). *Talented teenagers: The roots of success and failure.* Cambridge, England: Cambridge University Press.

Cuban, L. (1993). *How teachers taught: Constancy and change in American classrooms, 1890-1990.* New York: Teachers College Press.

Cummins, J. (1986). Empowering minority students: A framework for intervention. *Harvard Educational Review, 56*(1), 18-36.

Cunningham, D. (1992). Assessing constructions and constructing assessments. In T. Duffy & D. Jonassen (Eds.), *Constructivism and the technology of instruction: A conversation* (pp. 107-143). Hillsdale, NJ: Erlbaum.

Darling-Hammond, L. (1993). Reframing the school reform agenda: Developing capacity for school transformation. *Phi Delta Kappan, 74*(10), 752-761.

Darling-Hammond, L. (1994). Performance-based assessment and educational equity. *Harvard Educational Review, 64*(1), 5-30.

Delandshere, G., & Petrosky, A. R. (1994). Capturing teachers' knowledge: Performance assessment. *Educational Researcher, 23*(5), 11-18.

Derman-Sparks, L. (1989). *Anti-bias curriculum: Tools for empowering young children.* Washington, DC: NAEYC.

Desforges, C., Holden, C., & Hughes, M. (1994). Assessment at key stage one: Its effects on parents, teachers, and classroom practice. *Research Papers in Education, 9*(2), 133-158.

Doran, R. L., Boorman, J., Chan, F., & Hejaily, N. (1993). Alternative assessment of high school laboratory skills. *Journal of Research in Science Teaching, 30*(9), 1121-1131.

Doran, R. L., & Sentman, J. R. (1994). Special and special education: A science education perspective. *Remedial and Special Education, 15*(2), 128-133.

Duffy, G. G. (1993). Rethinking strategy instruction: Four teachers development and their low achievers' understanding. *The Elementary School Journal, 93*(3), 231-248.

Duffy, T. M., & Jonassen, D. H. (Eds.). (1992). *Constructivism and the technology of instruction: A conversation.* Hillsdale, NJ: Lawrence Erlbaum. (ERIC Document Reproduction Service No. ED 364 198)

Dunbar, S. B., Koretz, D. M., & Hoover, H. D. (1991). Quality control in the development and use of performance assessments. *Applied Measurement in Education, 4*(4), 289-303.

Edlind, E. P., & Haensly, P. A. (1985). Gifts of mentorship. *Gifted Child Quarterly, 29*(2), 55-60.

Egan, K., & Gardner, H. (1992). An exchange: The unschooled mind: How children think and how schools should teach. *Teachers College Press, 94*(2), 397-407.

Eisner, E. W. (1994). Commentary: Putting multiple intelligences in context: Some questions and observations. *Teachers College Record, 95*(4), 555-560.

Ellingson, M. K., Haeger, W., & Feldhusen, J. F. (1986). The Purdue mentor program: A university-based mentorship experience for g/c/t children. *Gifted Child Today, 9*(2), 2-5.

Feldhusen, J. F. (1995). Talent development: The new direction in gifted education. *Roeper Review, 18*(2), 92.

Finson, K. D., & Beaver, J. B. (1994). Performance assessment: Getting started. *Science Scope, 18*(1), 44-49.

Ford, D. (1994a). *Reversing underachievement among gifted Black students: Promising practices and programs.* New York: Teachers College Press.

Ford, D. (1994b). *The recruitment and retention of Black students in gifted programs* (RBDM 9406). Storrs, CT: University of Connecticut, The National Research Center on the Gifted and Talented.

Frasier, M. M. (1987). The identification of gifted Black students: Developing new perspectives. *Journal for the Education of the Gifted, 10*(3), 155-180.

Frasier, M. M. (1989). Identification of gifted Black students: Developing new perspectives. In C. J. Maker & S. W. Schiever (Eds.), *Critical issues in gifted education: Defensible programs for cultural and ethnic minorities* (Vol. III, pp. 213-226). Austin, TX: Pro-Ed.

Frasier, M. M. (1995, February). Multiple criteria: What? Why? How? *NCAGT Newsletter, 18*(6), 1, 14-17.

Frasier, M. M., & Garcia, J. H., & Passow, A. H. (1995). *A review of assessment issues in gifted education and their implications for identifying gifted minority students* (Research Monograph 95203). Storrs, CT: University of Connecticut, The National Research Center on the Gifted and Talented.

Frasier, M. M., & Passow, A. H. (1994). *Toward a new paradigm for identifying talent potential* (Research Monograph 94112). Storrs, CT: University of Connecticut, The National Research Center on the Gifted and Talented.

Frechtling, J. A. (1991). Performance assessment: Moonstruck or the real thing? *Educational Measurement: Issues and Practice, 10*(4), 23-25.

Freedman, M. (1989). Fostering intergenerational relationships for at-risk youth. *Gifted Child Today, 18*(2), 10-15.

Frey, D. E. (1991). Psychosocial needs of the gifted adolescent. In M. Bireley & J. Genshaft (Eds.), *Understanding the gifted adolescent* (pp. 35-49). New York: Teachers College Press.

Frisby, C. L., & Tucker, C. M. (1993). Black children's perception of self: Implications for educators. *The Educational Forum, 57*(2), 146-155.

Fullan, M. G. (1985). Change processes and strategies at the local level. *The Elementary School Journal, 85*(3), 390-421.

Fullan, M. G. (1991). *The new meaning of educational change.* New York: Teachers College Press.

Fullan, M. G. (1993a). *Teacher development and educational change.* Washington, DC: Falmer Press.

Fullan, M. G. (1993b). Why teachers must become change agents. *Educational Leadership, 50*(6), 12-17.

Gallagher, J. (1985). *Teaching the gifted child.* Boston: Allyn & Bacon.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences.* New York: Basic Books.

Gardner, H. (1984). Assessing intelligences: A comment on 'testing intelligence without I.Q. tests.' *Phi Delta Kappan, 65*(10), 699-700.

Gardner, H. (1988). Beyond the IQ: Education and human development. *National Forum, 68*(2), 4-7.

Gardner, H. (1991a). *The unschooled mind: How children think and how schools should teach.* New York: Basic Books.

Gardner, H. (1991b). Assessment in context: The alternative to standardized testing. In B. R. Gifford & M. C. O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement, and instruction* (pp. 77-120). Boston: Kluwer.

Gardner, H. (1993a). In a nutshell. In H. Gardner (Ed.), *Multiple intelligences* (pp. 5-12). New York: Basic Books.

Gardner, H. (1993b). *Multiple intelligences.* New York: Basic Books.

Gardner, H. (1994). Intelligences in theory and practice: A response to Elliot W. Eisner, Robert J. Sternberg, and Henry M. Levin. *Teachers College Record, 95*(4), 576-583.

Gardner, H., & Hatch, T. (1989). Multiple intelligences go to school: Educational implications of the theory of multiple intelligences. *Educational Researcher, 18*(8), 4-9.

Gardner, H., & Hatch, T. (1990). *Multiple intelligences go to school: Educational implications of the theory of multiple intelligences* (Technical Report No. 4). New York: Center for Technology in Education. (ERIC Document Reproduction Service No. ED 324 366)

Gear, G. H. (1976). Accuracy of teacher judgment in identifying intellectually gifted children: A review of the literature. *Gifted Child Quarterly, 20*(4), 478-490.

Gilbert, S. E. & Gay, G. (1989). Improving the success in school of poor Black children. In B. J. Robinson-Shade (Ed.), *Culture, style, and the educative process* (pp. 275-283). Springfield, IL: Charles C. Thomas.

Goertzel, M., Goertzel, V., & Goertzel, T. (1978). *300 eminent personalities.* San Francisco: Jossey-Bass.

Grant, C. A., & Sleeter, C. E. (1985). The literature on multicultural education: Review and analysis. *Educational Review, 37*(2), 97-118.

Greenwood, C. R. (Guest Ed.). (1994). Technology-based assessment within special education [Special Issue]. *Exceptional Children, 61*(2).

Grinder, R. E. (1985). The gifted in our midst: By their divine deeds, neurosis, and mental test scores we have known them. In F. D. Horowitz & M. O'Brien (Eds.), *The gifted and talented: Developmental perspectives* (pp. 5-36). Washington, DC: American Psychological Association.

Guild, G. P. (1993). Ethos for equality: A Massachusetts high school works toward harmony. *Teaching Tolerance, 2*(1), 23.

Guskey, T. R. (1994). What you assess may not be what you get. *Educational Leadership, 51*(6), 51-54.

Haertel, E. H. (1994). Theoretical and practical considerations. In T. R. Guskey (Ed.), *High stakes performance assessment: Perspectives on Kentucky's educational reform* (pp. 65-75). Thousand Oaks, CA: Corwin Press.

Hambleton, R. K., & Murphy, E. (1992). A psychometric perspective on authentic measurement. *Applied Measurement in Education, 5*(1), 1-16.

Hargreaves, A., & Fullan, M. (Eds.). (1992a). *Teacher development and educational change*. London: Falmer.

Hargreaves, A., & Fullan, M. (Eds.). (1992b). *Understanding teacher development.* New York: Teachers College Press.

Hatch, T., & Gardner, H. (1990). If Binet had looked beyond the classroom: The assessment of multiple intelligences. *International Journal of Educational Research, 14*(5), 415-429.

Hendricks, J., & Scott, M. (1987). Mentor companions in curiosity: A program for accepting and encouraging curiosity in young gifted children. *The Creative Child and Adult Quarterly, 12*(2), 11-123.

Hennessey, B. A., & Amabile, T. M. (1988). Story-telling: A method for assessing children's creativity. *Journal of Creative Behavior, 22*(4), 235-246.

Herman, J. L., Aschbacher, P. R., & Winters, L. (1992). *A practical guide to alternative assessment.* Alexandria, VA: Association for Supervision and Curriculum Development.

Herman, J. L., & Winters, L. (1994). Portfolio research: A slim collection. *Educational Leadership, 52*(2), 48-55.

High, M. H., & Udall, A. J. (1983). Teacher ratings of students in relation to ethnicity of students and school ethnic balance. *Journal for the Education of the Gifted, 6,* 154-165.

Hills, T. W. (1992). Reaching potentials through appropriate assessment. In S. Bredekamp & T. Rosegrant (Eds.), *Reaching potentials: Appropriate curriculum and assessment for young children* (Vol. 1, pp. 43-63). Washington, DC: National Association for the Education of Young Children.

Hollinger, C. L. (1991). Facilitating the career development of young gifted women. *Roeper Review, 13*(3), 135-139.

Huberman, A. M. (1981). *School-university collaboration supporting school improvement. Volume I: The Midwestern state teacher network case.* Washington, DC: American University. (ERIC Document Reproduction Service No. ED 210 260)

Huberman, A. M. (1983). School improvement strategies that work: Some scenarios. *Educational Leadership, 41*(3), 23-27.

Huberty, C. J. (1994). *Applied discriminant analysis.* New York: John Wiley & Sons.

Hunsaker, S. L., Frasier, M. M., Frank, E., Finley, V., & Klekota, P. (1995). *Performance of economically disadvantaged students placed in gifted programs through research-based assessment* (Research Monograph 95208). Storrs, CT: University of Connecticut, The National Research Center on the Gifted and Talented.

Instructional Objectives Exchange. (1972). *Measures of self-concept: Grades K-12* (rev. ed.). Los Angeles: Author.

Jackson, S. A. (1993). Opportunity to learn: The health connection. *Journal of Negro Education, 62*(3), 377-393.

Jonassen, D. (1990). Thinking technology: Toward a constructivist view of instructional design. *Educational Technology, 30*(9), 32-34.

Jonassen, D. (1991). Objectivism vs. constructivism: Do we need a philosophical paradigm shift? *Educational Technology, 31*(3), 5-14.

Jöreskeg, K. G., & Sörbom, D. (1989). *LISREL7: A guide to the program applications.* Chicago: SPSS, Inc.

Kanevsky, L. (1990). Pursuing qualitative differences in the flexible use of a problem-solving strategy by young children. *Journal for the Education of the Gifted, 13*(2), 115-140.

Karnes, M., & Johnson, L. (1986). Identification and assessment of gifted/talented handicapped and nonhandicapped children in early childhood. *Journal of Children in Contemporary Society, 18*(3-4), 35-54.

Kaufmann, F. A., Harrel, G., Milam, C. P., Woolverton, N., & Miller, J. (1986). The nature, role, and influence of mentors in the lives of gifted adults. *Journal of Counseling and Development, 64*(9), 576-578.

Kerr, B. A. (1985). Smart girls, gifted women: Special guidance concerns. *Roeper Review, 8*(1), 30-33.

Kitano, M. (1986). Evaluating program options for young gifted children. *Journal of Children in Contemporary Society, 18*(3-4), 89-101.

Knuth, R. A. & Cunningham, D. J. (1993). Tools for constructivism. In T. Duffy, J. Lowyck, & D. Jonassen (Eds.), *Designing environments for constructive learning* (pp. 163-189). Heidelberg, Germany: Springer-Verlag.

Kornhaber, M., & Gardner, H. (1993). *Varieties of excellence: Identifying and assessing children's talents.* New York: The National Center for Restructuring Education, Schools, and Teaching.

Kotulak, R. (1996). *Inside the brain: Revolutionary discoveries of how the mind works.* Kansas City, MO: Andrews and McMeel.

Krechevsky, M. (1991). Project Spectrum: An innovative assessment alternative. *Educational Leadership, 48*(5), 43-48.

Lane, S. (1993). The conceptual framework for the development of a mathematics performance assessment instrument. *Educational Measurement: Issues and Practices, 12*(2), 16-23.

Lane, S., Stone, C. A., Ankenmann, R. D., & Liu, M. (1994). Reliability and validity of a mathematics performance assessment. *International Journal of Educational Research, 21*(3), 247-266.

Larkin, J. M. (1993). Rethinking basic skills instruction with urban students. *Educational Forum, 57*(4), 413-419.

Lawrenz, F. (1992). Authentic assessment. In F. Lawrenz, K. Cochran, J. Krajcik, & P. Simpson (Eds.), *Research matters . . . to the science teacher* (NARST Monograph No. 5, pp. 65-70). Manhattan, KS: National Association of Research in Science Teaching.

Linder, C. J. (1993). A challenge to conceptual change. *Science Education, 77*(3), 292-300.

Linn, R. L. (1994). Performance assessment: Policy promises and technical measurement standards. *Educational Researcher, 23*(9), 4-14.

Linn, R. L., Baker, E., & Dunbar, S. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher, 20*(8), 15-21.

Linn, R. L., & Burton, E. (1994). Performance-based assessment: Implications of task specificity. *Educational Measurement: Issues and Practices, 13*(1), 5-8, 15.

Lupkowski, A. E., Assouline, S. G., & Vestal, J. (1992). Mentors in math. *Gifted Child Today, 15*(3), 26-31.

MacGinitie, W. H. (1993). Some limits of assessment. *Journal of Reading, 36*(7), 556-560.

Maeroff, G. I. (1991). Assessing alternative assessment. *Phi Delta Kappan, 73*(4), 273-281.

Maker, C. J. (1993). Creativity, intelligence, and problem solving: A definition and design for cross-cultural research and measurement related to giftedness. *Gifted Education International, 9,* 68-77.

Maker, C. J., Nielson, A. B., & Rogers, J. A. (1994). Giftedness, diversity, and problem-solving. *Teaching Exceptional Children, 27*(1), 4-18.

Maker, C. J., Rogers, J. A., & Nielson, A. B. (1992). *Assessment/observation of problem solving abilities.* Tucson, AZ: University of Arizona.

Maker, C. J., & Schiever, S. W. (Eds.). (1989). *Critical issues in gifted education: Defensible programs for cultural and ethnic minorities, Vol. II.* Austin, TX: Pro-ed.

Marzano, R. J. (1994). Lessons from the field about outcome-based performance assessments. *Educational Leadership, 51*(6), 44-50.

Matthews, D. (1988). Gardner's multiple intelligence theory: An evaluation of relevant research literature and a consideration of its application to gifted education. *Roeper Review, 11*(2), 100-104.

Maturana, H. R., & Varela, F. J. (1987). *The tree of knowledge: The biological roots of human understanding.* Boston: New Science Library.

McArdle, J. J., & McDonald, R. P. (1984). Some algebraic properties of the reticular action model for moment structures. *British Journal of Mathematical and Statistical Psychology, 37*(2), 234-251.

McPartland, J. M., & Nettles, S. M. (1991). Using community adults as advocates or mentors for at-risk middle school students: A two-year evaluation of Project RAISE. *American Journal of Education, 99*(4), 568-585.

Merriam, S. (1983). Mentors and protégés: A critical review of the literature. *Adult Education Quarterly, 30*(3), 161-173.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher, 23*(2), 13-23.

Meyer, C. A. (1992). What's the difference between "authentic" and "performance" assessment? *Educational Leadership, 49*(8), 39-40.

Miller, M. D., & Legg, S. M. (1993). Alternative assessment in a high-stakes environment. *Educational Measurement: Issues and Practices, 12*(2), 9-15.

Milman, C. P., & Schwartz, B. (1992). The mentorship connection. *Gifted Child Today, 15*(3), 9-13.

Minnehan, J. S., & Strunk, B. S. (1992). Academic mentoring in a secondary school. *Gifted Child Today, 15*(3), 43-45.

Moon, T. R. (1993). *A teacher's guide to developing performance-based assessments.* Charlottesville, VA: Bureau of Educational Research, University of Virginia.

Murrell, P. (1993). Afrocentric immersion: Academic and personal development of African American males in public schools. In T. Perry & J. W. Fraser (Eds.), *Freedom's plow: Teaching in the multicultural classroom* (pp. 231-259). New York: Routledge.

NAEYC. (1988). Position statement on standardized testing of young children 3 through 8 years of age. *Young Children, 43*(3), 42-47.

Nuttall, D. L. (1992). Performance assessment: The message from England. *Educational Leadership, 49*(8), 54-57.

Ogbu, J. U. (1988). Human intelligence testing: A cultural-ecological perspective. *National Forum, 68*(2), 23-29.

O'Neil, J. (1992). Putting performance assessment to the test. *Educational Leadership, 49*(8), 14-19.

Papert, S., & Harel, I. (1990). Software design as a learning environment. *Interactive Learning Environment, 1*(1), 1-32.

Patton, J. M. (1992). Assessment and identification of African-American learners with gifts and talents. *Exceptional Children, 59*(2), 150-159.

Patton, J. M., Prillaman, D., & VanTassel-Baska, J. (1990). The nature and extent of programs for the disadvantaged/gifted in the United States and territories. *Gifted Child Quarterly, 34*(3), 94-96.

Peers, M. G. (1993). A teacher/researcher's experience with performance-based assessment as a diagnostic tool. *Journal of Reading, 36*(7), 544-548.

Perry, T., & Fraser, J. W. (1993). Reconstructing schools as multiracial/multicultural democracies. In T. Perry & J. W. Fraser (Eds.), *Freedom's plow: Teaching in the multicultural classroom* (pp. 231-259). New York: Routledge.

Piirto, J. (1994). *Talented children and adults: Their development and education.* Columbus, OH: Merrill/Prentice-Hall.

Pink, W. T. (1986). Facilitating change at the school level: A missing factor in school reform. *Urban Review, 18*(1), 19-30.

Pizzini, E. L. (1985). Improving science for instruction for gifted high school students. *Roeper Review, 7*(4), 231-234.

Plewis, J. (1991). Pupils' progress in reading and mathematics during primary school: Associations with ethnic group and sex. *Educational Research, 33*(2), 133-140.

Poteet, J. A., Choate, J. S., & Stewart, S. C. (1993). Performance assessment and special education: Practices and prospects. *Focus on Exceptional Children, 26*(1), 1-20.

Prather, J. P. (1993). Reform revisited: The trend toward constructivist learning. *Journal of Elementary Science Education, 5*(2), 52-70.

Puckett, M. B., & Black, J. K. (1994). *Authentic assessment of the young child.* New York: Merrill.

Reckase, M. (Ed.). (1993). Performance assessment [Special Issue]. *Journal of Educational Measurement, 30*(3).

Reichel, A. G. (1994). Performance assessment: Five practical approaches. *Science and Children, 32*(2), 21-25.

Renzulli, J. S. (1973). Talent potential in minority group students. *Exceptional Children, 39*(4), 437-444.

Renzulli, J. S. (1977). *The enrichment triad: A guide for developing defensible programs for the gifted.* Mansfield, CT: Creative Learning Press.

Renzulli, J. S. (1986). The three-ring conception of giftedness: A developmental model for creative productivity. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness,* (pp. 53-92). Cambridge, MA: Cambridge University Press.

Renzulli, J. S. (1994). *Schools for talent development: A practical plan for total school improvement.* Mansfield Center, CT: Creative Learning Press.

Renzulli, J. S., Reis, S. M., & Smith, L. H. (1981). *The revolving door identification model.* Mansfield, CT: Creative Learning Press.

Richert, E. S. (1987). Rampant problems and promising practices in the identification of disadvantaged gifted students. *Gifted Child Quarterly, 31*(4), 149-154.

Richert, E. S. (1997). Excellence with equity in identification and programming. In N. Colangelo & G. A. Davis (eds.), *Handbook of gifted education* (2nd. ed., pp. 75-88). Boston: Allyn & Bacon.

Riley, M. K. (1993). Walking the talk: Putting constructivist thinking into practice in classrooms. *Educational Horizons, 71*(4), 187-196.

Rivera, D. (1993). Performance, authentic, and portfolio assessment: Emerging alternative assessment options in search of an empirical basis. *Diagnostique, 18*(4), 325-348.

Robinson, A. (1994). Assessment, identification, and evaluation. *Tempo, 14*(2), 1, 22-23.

Roche, G. (1979). Much ado about mentors. *Harvard Business Review, 57,* 2014-16, 2020, 2024; 2026-28.

Rosenthal, R., & Rubin, D. B. (1979). A note on percent variance explained as a measure of the importance of effects. *Journal of Applied Social Psychology, 9*(5), 395-396.

Rosnow, R. L., & Rosenthal, R. (1988). Focused tests of significance and effect size estimation in counseling psychology. *Journal of Counseling Psychology, 35*(2), 203-208.

Sarason, S. (1990). *The predictable failure of educational reform.* San Francisco: Jossey-Bass.

Semple, B. M. (1992). *Performance assessment: An international experiment* (ETS Report No. 22-CAEP-06). Princeton, NJ: IAEP/ETS.

Shade, B. (1981). *Afro-American style: A variable in school success?* (Report WRDCIS-TP-94). Washington, DC: National Institute of Health. (ERIC Document Reproduction Service No. ED 211 657)

Shade, B. (1989). *Culture, style and the educative process.* Springfield, IL: Thompson.

Shade, B. (1994). Understanding the African-American learner. In E. R. Hollins, J. E. King, & W. C. Hayman (Ed.), *Teaching diverse populations: Formulating a knowledge base* (pp. 175-189). New York: Holt, Rinehart, and Winston.

Shaklee, B. (1992). Identification of young gifted students. *Journal for the Education of the Gifted, 15*(2), 134-144.

Shavelson, R. J., Baxter, G. P., & Pine, J. (1991). Performance assessment in science. *Applied Measurement in Science, 4*(4), 347-362.

Showers, B. (1985). Teachers coaching teachers. *Educational Leadership, 42*(7), 43-48.

Simons, P. R. (1993). Constructive learning: The role of the learner. In T. Duffy, J. Lowyck, & D. Jonassen (Eds.), *Designing environments for constructive learning* (pp. 291-315). Heidelberg, Germany: Springer-Verlag.

Sleeter, C. E., & Grant, C. A. (1987). An analysis of multicultural education in the United States. *Harvard Educational Review, 57*(4), 421-444.

Sleeter, C. E., & Grant, C. A. (1994). *Making choices for multicultural education: Five approaches to race, class, and gender.* New York: Macmillan.

Slicker, E. K., & Palmer, D. J. (1993). Mentoring at-risk high school students: Evaluation of a school-based program. *The School Counselor, 40*(5), 327-334.

Smith, G. P., Miller, M. C., & Joy, J. (1988). A case study of the impact of performance-based testing on the supply of minority teachers. *Journal of Teacher Education, 33*(4), 45-53.

Spradley, J. (1979). *The ethnographic interview.* New York: Holt, Rinehart and Winston.

Sternberg, R. J. (1984). Fighting butter battles. *Phi Delta Kappan, 65*(10), 700.

Sternberg, R. J. (1986). A triarchic theory of intellectual giftedness. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 223-243). New York: Cambridge University Press.

Sternberg, R. J. (1988). Beyond IQ testing. *National Forum, 68*(2), 8-11.

Stiggins, R. J. (1991). Facing the challenges of a new era of educational assessment. *Applied Measurement in Education, 4*(4), 263-273.

Stiggins, R. J., & Plake, B. (Eds.). (1991). Performance assessment [Special issue]. *Applied Measurement in Education, 4*(4).

Swezey, R. W. (1981). *Individual performance assessment: An approach to criterion-referenced test development.* Reston, VA: Reston Publishing.

Tabachnik, B. G., & Fidell, L. S. (1989). *Multivariate statistics* (2nd ed.). New York: Harper Collins.

Tannenbaum, A. J. (1983). *Gifted children: Psychological and educational perspectives.* New York: Macmillan.

Tharp, R. (1989). Psychosocial variables and constants: Effects on teaching and learning in schools. *American Psychologist, 44*(2), 349-359.

Thorndike, R. L., & Hagen, E. (1955). *Measurement and evaluation in psychology and education.* New York: Wiley & Sons.

Tizard, J., Schofield, W. N., & Hewison, J. (1982). Collaboration between teachers and parents in assisting children's reading. *British Journal of Educational Psychology, 52*(1), 1-15.

Torrance, E. P. (1984). *Mentor relationships: How they aid creative achievement, endure, change, and die.* New York: Bearly Limited.

Udall, A. J., & Passe, M. (1993). *Gardner-based/performance-based assessment notebook.* Charlotte, NC: Charlotte-Mecklenburg Schools.

U.S. Department of Education. (1993). *National excellence: A case for developing America's talent.* Washington, DC: Office of Educational Research and Improvement.

Vadasy, P., & Maddox, M. (1992). *Yakima equity study: The conditions of success of migrant, Hispanic, and Native American students in the Yakima Valley.* Yakima, WA: Washington Research Institute.

Vann, K. R., & Kunjufu, J. (1993). The importance of an Afrocentric, multicultural curriculum. *Phi Delta Kappan, 74*(6), 490-491.

Verma, G., & Mallick, K. (1988). Self-esteem and educational achievement in British young South Asians. In G. Verma & P. Pumfry (Eds.), *Educational attainments: Issues and outcomes in multicultural education* (pp. 152-164). London: Falmer Press.

Walker-Dalhouse, D. (1993). Beginning reading and the African-American child at risk. *Young Children, 49*(1), 24-28.

Wallace, M. (1991). Flexible planning: A key to the management of multiple innovations. *Educational Management and Administration, 19*(3), 180-192.

Walters, J. (1992). Application of multiple intelligences: Research in alternative assessment. *Focus on evaluation and measurement*, Volumes 1 and 2. Washington, DC: Proceedings of the National Symposium on Limited English Proficient Student Issues. (ERIC Document Reproduction Service No. ED 349 812)

Walters, J. M., & Gardner, H. (1985). The development and education of intelligences. In F. Link (Ed.), *Essays on the intellect* (pp. 1-21). Washington, DC: Curriculum Development Associates.

Walters, J., & Gardner, H. (1986). The theory of multiple intelligences: Some issues and answers. In R. J. Sternberg & R. Wagner (Eds.), *Practical intelligences* (pp. 163-181). New York: Cambridge University Press.

Weinberg, R. A. (1989). Intelligence and IQ: Landmark issues and great debates. *American Psychologist, 44*(2), 98-104.

West, A., Sammons, P., Hailes, J., & Nuttall, D. (1994). The standard assessment tasks and the boycott at key stage 1: Teachers' and head teachers' views in six inner-city schools. *Research Papers in Education, 9*(3), 321-337.

Wiggins, G. (1993). Assessment: Authenticity, context, and validity. *Phi Delta Kappan, 75*(3), 200-214.

Willis, S. (1994). Teaching language-minority students. *ASCD Update, 36*(5), 1-5.

Wolf, D., Bixby, J., Glenn III, J., & Gardner, H. (1991). To use their minds well: Investigating new forms of student assessment. In G. Grant (Ed.), *Review of research in education: Volume 17* (pp. 31-74). Washington, DC: American Educational Research Association.

Wolf, D. P., LeMahieu, P. G., & Eresh, J. (1992). Good measure: Assessment as a tool for educational reform. *Educational Leadership, 49*(8), 8-13.

Wong, S. C. (1993). Promises, pitfalls, and principles of text selection in curricular diversification: The Asian-American case. In T. Perry & J. W. Fraser (Eds.), *Freedom's plow: Teaching in the multicultural classroom* (pp. 47-65). New York: Routledge.

Worthen, B. R. (1993). Critical issues that will determine the future of alternative assessment. *Phi Delta Kappan, 74*(6), 444-454.

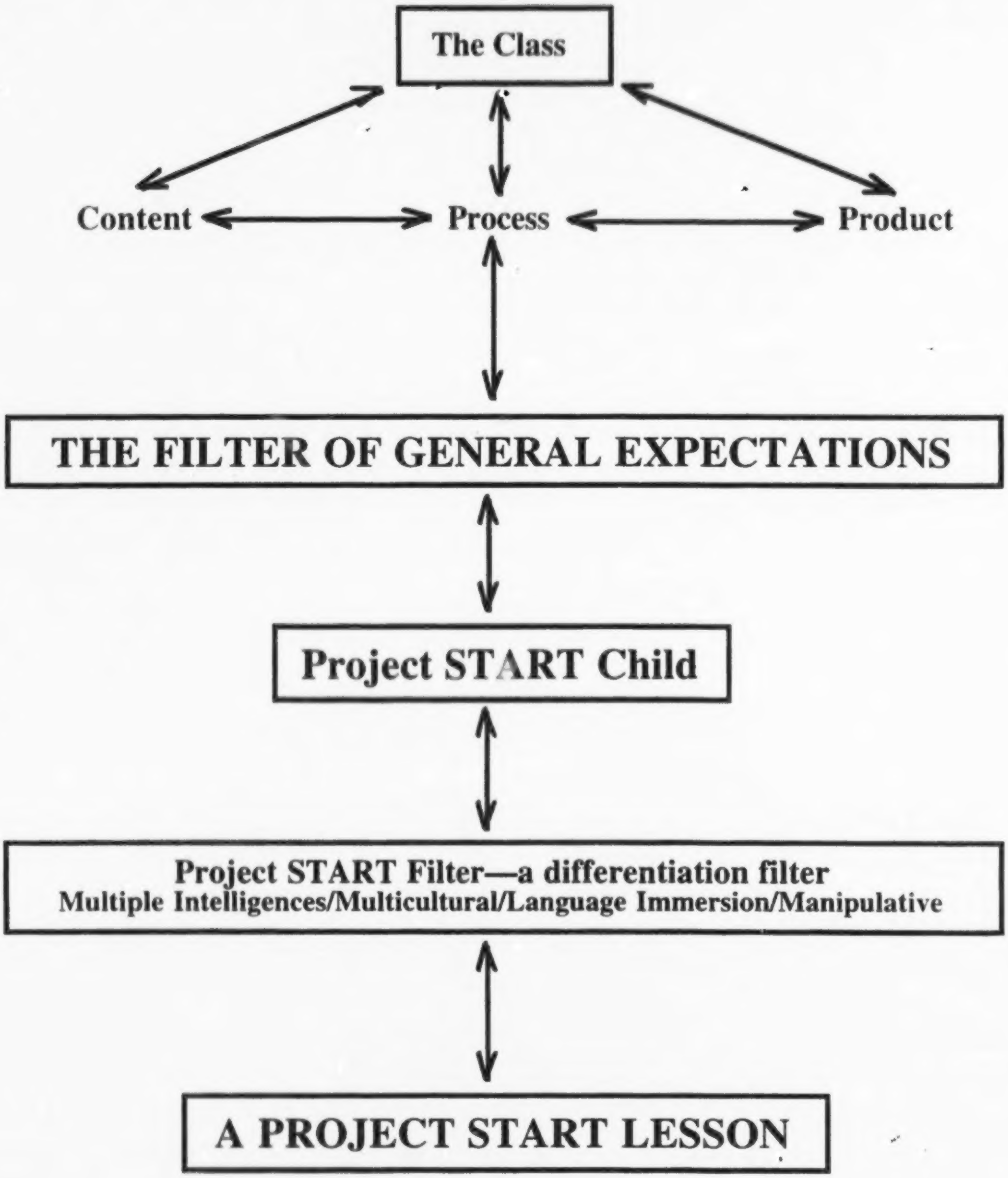
Wright, L., & Borland, J. H. (1992). A special friend: Adolescent mentors for young, economically disadvantaged, potentially gifted students. *Roeper Review, 14*(3), 124-129.

Zuck, J. C. (1983). Global citizenship, pluralistic societies, and teacher education. In L. V. D. Berg-Eldering, F. J. M. De Rijcke, & L. V. Zuck (Eds.), *Multicultural education: A challenge for teachers* (pp. 173-191). Dordrecht, Holland: Foris.

Appendix A

Draft Model of Developing a Project START Lesson

DRAFT MODEL OF DEVELOPING A PROJECT START LESSON


The Class
Content
Process
Product
THE FILTER OF GENERAL EXPECTATIONS
Project START Child
Project START Filter—a differentiation filter
Multiple Intelligences/Multicultural/Language Immersion/Manipulative
A PROJECT START LESSON

Appendix B

START Classroom Observation Protocols

Classroom Practices Record (CPR)
Form VA2

CPR: Identification Information

Date of Observation_____ Observer__________________ Length of Observation______

START Teacher________________________ School____________________________

Grade Level_______ No. of START Students in Classroom_______

No. of START Girls_______ No. of START Boys_______

Total Number of Students in the Classroom _____ No. of Girls_____ No. of Boys _____

Subject(s) Observed

_____ Math
_____ Language Arts
_____ Other (list)____
_____ Social Studies
_____ Science

Abilities on Which Target Students are Identified (record # for each category)

_____ Spatial
_____ Verbal/Linguistic
_____ Logical/Mathematical
_____ Personal

CPR: Physical Environment Inventory

During each observation, please place an "x" in the appropriate blank to reflect working arrangements observed:

Learning Centers/Work Groups

_____ 3 or more learning/interest centers
_____ 2 learning/interest centers
_____ 1 learning/interest center
_____ No learning/interest centers
_____ 3 or more small working groups
_____ 2 small working groups
_____ 1 small working group
_____ No small working groupss

CPR Instrument Developed by the University of Connecticut—The National Research Center on the Gifted and Talented
(Forms VA & VA2 Modified by the University of Virginia Site—The National Research Center on the Gifted and Talented)

Teacher Observed____________________________

Project START Classroom Observation Checklist

Instructional Characteristics & Descriptors	Observer Comments
1. Project START Components Use of manipulatives________ Evidence of multicultural emphasis________ Language immersion________ Varied intelligence options for learning, expressing learning and/or assessing learning____________	
2. Diffrentiated Learning Environment Assessment of student readiness for task________ Activities/materials/assessments for students at different levels of readiness________ Activities/assessments which allow for differing learning styles________ Teacher receptivity to diverse learners and learner needs________ Degree of student-centeredness reflected through discussion, sense-making, decision-making, and "time on center stage"___________	
3. Content of Instruction Balance of skills and key ideas (concepts)________ Meaningful application of skills and concepts by students to solve problems and/or create products to be shared with others________ Balance of content and process________ Evidence of critical and creative thinking________	

Key: 1=little evidence 3=some evidence 5=strong evidence
n/a=not applicable in this observation

Observer:________________________________ Date of Observation:_____________

Research Monograph Series

The National Research Center on the Gifted and Talented
The University of Connecticut
362 Fairfield Road, U-7
Storrs, CT 06269-2007
www.gifted.uconn.edu

Production Assistants
Dawn R. Guenther
Siamak Vahidi

Reviewers
Beverly Coleman
Gary Davis
Bessie Duncan
E. Jean Gubbins
Leticia Hernandez
Valentina Kloosterman
Harry Milne
Emily Stewart

Also of Interest

Regular Classroom Practices With Gifted Students: Results of a National Survey of Classroom Teachers
Francis X. Archambault, Jr., et al.

An Observational Study of Instructional and Curricular Practices Used With Gifted and Talented Students in Regular Classrooms
Karen L. Westberg, et al.

Why Not Let High Ability Students Start School in January?
The Curriculum Compacting Study
Sally M. Reis, et al.

Evaluation of the Effects of Programming Arrangements on Student Learning Outcomes
Marcia A. B. Delcourt, et al.

Qualitative Extension of the Learning Outcomes Study
Marcia A. B. Delcourt and Karen Evans

Also of interest from the

Research Monograph Series

Toward a New Paradigm for Identifying Talent Potential
Mary M. Frasier and A. Harry Passow

Talents in Two Places: Case Studies of High Ability Students With Learning Disabilities Who Have Achieved
Sally M. Reis, Terry W. Neu, and Joan M. McGuire

The Effects of Group Composition on Gifted and Non-Gifted Elementary Students in Cooperative Learning Groups
David A. Kenny, Francis X. Archambault, Jr., and Bryan W. Hallmark

Extending the Pedagogy of Gifted Education to All Students
Sally M. Reis, Marcia Gentry, and Sunghee Park

Case Studies of Talented Students Who Achieve and Underachieve in an Urban High School
Sally M. Reis, et al.

Profiles of Successful Practices for High Ability Students in Elementary Classrooms
Karen L. Westberg and Francis X. Archambault, Jr.

Academic Diversity in the Middle School: Results of a National Survey of Middle School Administrators and Teachers
Tonya R. Moon, Carol A. Tomlinson, and Carolyn M. Callahan

A Triarchic Approach to Giftedness
Robert J. Sternberg

A Study of Achievement and Underachievement Among Gifted, Potentially Gifted, and Average African-American Students
Donna Y. Ford

Instruments Used in the Identification of Gifted and Talented Students
Carolyn M. Callahan, et al.

Instruments and Evaluation Designs Used in Gifted Programs
Carolyn M. Callahan, et al.

Preservice Teacher Preparation in Meeting the Needs of Gifted and Other Academically Diverse Students
Carol A. Tomlinson, et al.

NRC
G/T

The National Research Center on the Gifted and Talented Research Teams

The University of Connecticut

Dr. Francis X. Archambault, Associate Director
The University of Connecticut
School of Education, U-4
Storrs, CT 06269-2004
860-486-4531

Dr. Alexinia Y. Baldwin
Dr. Scott W. Brown
Dr. Deborah E. Burns
Dr. David A. Kenny
Dr. Jonna Kulikowich
Dr. Sally M. Reis
Dr. Karen L. Westberg
Dr. Michael F. Young

The University of Georgia

Dr. Mary M. Frasier, Associate Director
The University of Georgia
Department of Educational Psychology
323 Aderhold Hall
Athens, GA 30602-7146
404-542-5106

Dr. Scott L. Hunsaker

The University of Virginia

Dr. Carolyn M. Callahan, Associate Director
Curry School of Education
The University of Virginia
405 Emmet Street
Charlottesville, VA 22903
804-982-2849

Dr. Michael S. Caldwell
Dr. Marcia A. B. Delcourt
Dr. Kathleen May
Dr. Claudia Sowa
Dr. Ellen Tomchin
Dr. Carol A. Tomlinson

Yale University

Dr. Robert J. Sternberg, Associate Director
Department of Psychology
Yale University
P.O. Box 208205
New Haven, CT 06520-8205
203-432-4632

Dr. Pamela Clinkenbeard